FUNDAMENTALS
OF
QUANTUM
MECHANICS

V. A. FOCK

Fundamentals of Quantum Mechanics

by Vladimir A. Fock,
Member of USSR Acad. Sci.

This unique course in quantum mechanics was written by an outstanding physics theoretician, who contributed much to the founding and development of the quantum theory. Many of the chapters bear the stamp of his scientific investigations.

The first Russian edition appeared in 1932 and for some years it was the only Soviet book on the quantum theory. It has retained its significance to this day, but has long been a bibliographical rarity.

For the second Russian edition the author has revised and greatly augmented the book, including his latest findings in quantum mechanics. He elaborates on the epistomological bases of the science, devoting several sections to specific problems that add to an understanding of the theory. Two of the new chapters consider Pauli's theory of the electron and the many-electron problem in its application to the atomic theory.

ИЗДАТЕЛЬСТВО • МИР •

В. А. ФОК

НАЧАЛА КВАНТОВОЙ МЕХАНИКИ

V. A. FOCK

FUNDAMENTALS

OF

QUANTUM MECHANICS

Translated from the Russian
by
Eugene Yankovsky

MIR PUBLISHERS
MOSCOW

FOREWORD

Vladimir Aleksandrovich Fock was one of the group of brilliant physics theoreticians whose work built the magnificent edifice of the quantum theory.

Forty five years have passed since he wrote the first edition of this unique course in quantum mechanics, which was the first of its kind in the Soviet Union and one of the first in the world.

Despite the 45 years, the nature of this book, the organization of the material, and the interpretation of specific topics all appear more natural today than when the book was first published. This is a feature of none but books whose authors have taken an active part in creating the subject of research and have understood it much more deeply than most of their contemporaries.

Practically every section of the course has to do with the author's own research, much of which has become an essential element of the quantum theory. This lends the course the special attraction of being basic.

As the title indicates, the book lays no claim to complete coverage of the subject. It gives only the simplest applications in addition to the basic concepts of the theory. Many of the more complex problems, such as the theories of molecules, atomic nuclei, and solids, are not included in the course.

The changes and additions to be found in the second edition are primarily related to the introductory part of the course, Chapter I, and concern the philosophical foundations of the quantum theory. Fock considered it of prime importance to formulate the basic concepts of quantum mechanics from the proper materialistic standpoint. The views outlined in Chapter I evolved from numerous discussions, some with Niels Bohr. In fact, there is evidence that Fock's criticism of "non-observability in principle" prompted Bohr to abandon this idea in his later works. Because of the attention paid to the epistemological aspects of the theory and their detailed and consistently materialistic interpretation, this book differs favourably from other courses.

The new edition is also augmented by sections devoted to the method of a self-consistent field, to the intrinsic symmetry of the hydrogen atom, and to other problems. Due to this it more fully reflects Fock's contribution to the quantum theory.

Finally, some revisions have been made in the section on the **Dirac equation and the positron theory**, which was such a sensa-

5

tion when the book was first written. Fock recognized the great-
ness of the theoretical prediction of the existence of antiparticles
but always emphasized the incompleteness of the positron theory,
since it is impossible to give an exact description of the processes
of creation and annihilation with a single-particle equation.

One of the remarkable features of Fock's scientific work is his
amazing mathematical powers, his ability to solve complicated
mathematical problems using the simplest and most unexpected
methods. This quality is especially evident in his scientific papers,
of course, but in this book too one can easily feel the author's
brilliant mathematical individuality combined with a precise and
simple manner of presentation.

The preparation of the second edition proved to be the last
work of V. A. Fock. We feel sure that those who use this book to
study quantum mechanics will experience the joy of dealing with
a primary source and feel the spirit of that wonderful time when
the horizons of human knowledge were immeasurably expanded
in a matter of years. Fock witnessed and participated in this
process.

The author died before the book came out, but he did succeed
in preparing most of it for press.

The Institute of Physics at   *Professor M. G. Veselov*
Leningrad State University   *Professor Yu. N. Demkov*
1976

PREFACE TO THE SECOND RUSSIAN EDITION

The second edition of this book, unlike the first, devotes a separate chapter to the nonrelativistic theory of the electron spin (Pauli's theory of the electron) and contains a chapter on the many-electron problem of quantum mechanics. In addition, some of the author's findings have been incorporated as separate sections. Otherwise, the subject matter of the book (both the mathematical theory and its physical interpretation) remains the same, except for certain new formulations of an epistemological character (the concepts of relativity with respect to the means of observation and of potential possibility), which has necessitated changing the expression "the statistical interpretation of quantum mechanics" to "the probabilistic interpretation". The new formulations are more precise than the previous ones.

The title of the book speaks for itself. The word "fundamentals" can be understood as "basic principles" or as "introductory facts".

We hope that although more than 40 years have passed since the book was written, the material in it has not lost its timeliness and will be useful to students of quantum mechanics.

1974   *V. A. Fock*

PREFACE TO THE FIRST RUSSIAN EDITION

This book was conceived as an elaboration of the reports on Dirac's theory of the electron delivered by the author at the Leningrad State Optical Institute in early 1929. The original plan was, however, expanded to include Dirac's theory, which is considered in Part III, the basic concepts of quantum mechanics (Part I), and Schrödinger's theory (Part II).

From the vast subject of the quantum theory the author has chosen material limited in two respects. First, the book considers none but the main principles and simplest applications of quantum mechanics. It concerns itself exclusively with the one-body problem. It does not deal with the many-body problem or the Pauli exclusion principle, basic to that problem. Second, the author has sought to confine himself to that part of the theory that is considered proved, that is, quantum mechanics proper. He has not examined quantum electrodynamics since this theory has yet to be fully elaborated.

The author's main purpose is to introduce the reader to a new set of ideas differing greatly from the classical theory. He has endeavoured to avoid using images from the classical theory as being inapplicable to quantum physics. Rather, he has attempted to familiarize the reader with the basic concepts underlying a quantum description of the states of atomic systems.

As for the presentation of material the author believes that a fairly detailed examination of the mathematics of a problem facilitates rather than hinders understanding, since it eliminates difficulties that the reader may encounter in dealing with the mathematical aspects and thus allows attention to be focussed on the physics of the problem.

This book is intended for senior students of physics and mathematics and persons with a sufficient preparation in mathematics.

Leningrad
The State Optical Institute
August 1931

V. A. Fock

CONTENTS

PART II

SCHRÖDINGER'S THEORY

PART V

DIRAC'S THEORY OF THE ELECTRON

BASIC CONCEPTS
OF QUANTUM MECHANICS

Chapter I

THE PHYSICAL AND EPISTEMOLOGICAL BASES
OF QUANTUM MECHANICS

1. The need for new methods and concepts in describing atomic phenomena

Quantum mechanics appeared during the first few decades of the century on the basis of studies of atomic phenomena. The structure of the atom, the properties of electrons and atomic nuclei, the very stability of a system consisting of a positively charged nucleus and negatively charged electrons, the radiation of light by atoms and molecules, and, last, the diffraction of electrons — all these properties and phenomena require for their explanation ideas and physical concepts that differ substantially from the ideas and concepts of classical physics.

A precise formulation of the new concepts demands new mathematical tools, and we will familiarize ourselves with these in subsequent chapters. But we will try to explain the principal difference between quantum mechanics and classical mechanics in this introductory chapter.

2. The classical description of phenomena

When we describe various phenomena in terms of classical physics, we assume that physical processes are independent of the conditions of observation. Thus we take it for granted that we can always "spy" on the process and yet not interfere with it or influence it. True enough, if we "spy" on a physical process from different view points (and correspondingly use different frames of reference for its description), it will appear to us in different ways. For instance, the free fall of a body may proceed in a straight line in one frame of reference and in a parabola in another. But the dependence of the form of a phenomenon on the frame of reference has always been taken into consideration, namely, by transforming from the coordinates of one frame of reference to the coordinates of another. Such a change in form introduces nothing new into the phenomenon. For this reason in

classical physics we can speak of the independence of a pheno-
menon from the manner of observation.

Quantum mechanics has shown that in microprocesses this is
not the case. Here the very possibility of observation presupposes
definite physical conditions that may be related to the essence of
the process. Specifying these conditions does not mean simply
indicating a particular frame of reference but requires more de-
tailed elaboration.

Neglect of these considerations leads to an abstraction that we
may call the *absolutization of physical processes*. If we accept
this abstraction, it becomes possible to consider physical processes
as occurring by themselves regardless of whether there is a real
possibility of their observation (that is, whether the appropriate
physical conditions exist for such processes).

The use of this abstraction is justified in studying macroscopic
phenomena, for in these the influence produced by a measurement
is to all practical purposes negligible. The absolutization of such
phenomena seemed so natural that before the appearance of
quantum mechanics it was never specifically stated. It went with-
out saying that physical processes occur by themselves, which
considerably simplified their description since there was no need
to specify the conditions of observation. All of classical physics is
based on the absolutization of physical processes. This abstraction
is one of its characteristic features.

Another abstraction permitted in classical physics is the
possibility of unlimited amendment of observation. By this we
mean not only an increasingly precise measurement of a specific
quantity but simultaneously the measurement of any other quan-
tity related to the observed object or phenomenon. This can be
called the *particularization of measurements*. Even when measur-
ing different quantities requires different conditions of observation,
classical physics considers it possible to combine the results
in an overall picture describing the physical process under
investigation. There is a logical connection between allow-
ing for the independence of the physical process from the con-
ditions of observation, that is, absolutization of the process, and
allowing for the possibility of encompassing different aspects
and characteristics of the behaviour of an object in the physical
process.

The concepts of classical physics prompt the idea that not only
an absolute but an exhaustive description of the state of motion
of a physical system (with certain degrees of freedom) is possible.
And an exhaustive description is assumed to be achieved if there
is full particularization of observations and further observations
can add nothing new.

3. Range of application of the classical way
of describing phenomena.
Heisenberg's and Bohr's uncertainty relations

Such fundamental facts as the wave-corpuscular duality of light and of particles of matter prove that the classical way of describing phenomena is unsuitable for micro-objects. At the same time we cannot dismiss it completely, since to describe phenomena objectively we must rely, directly or indirectly, on something that does not require reservations concerning the manner of observation. And this is the case with the "absolute" manner of description used in classical physics.

To apply the classical (absolute) manner of description intelligently we must first establish its limits. If we assume that the mathematical apparatus of quantum mechanics is known, the relations of classical physics derive from it in the form of a certain approximation, and the limits of application of the classical manner of description prove to be the conditions of applicability of this approximation. But in our discourse we proceed from classical mechanics and can only use the simplest quantum relationships.

Let us consider a simple case, the motion of a mass point with a mass m. In classical mechanics the state of motion of a mass point at any given moment of time is determined by its position (x, y, z) and momentum (p_x, p_y, p_z). It would be incorrect, however, to consider the two sets simultaneously without referring to the possibility of their measurement, which is limited by quantum effects.

As Werner Heisenberg proved, the localization of a particle in space demands conditions that are not favourable for measuring its momentum, that is, for the localization of the particle in momentum space. Conversely, conditions that are needed to measure the momentum of a particle preclude the possibility of localizing the particle in ordinary space.

Quantum effects, which limit the possibility of measurement, manifest themselves, for instance, when light quanta irradiate a particle. What is important here is that a photon, which is characterized by wave parameters, is at the same time a bearer of definite energy and momentum, which makes it a "particle of light". The wave parameters are: the frequency ν (or the angular frequency $\omega = 2\pi\nu$), the wavelength $\lambda = c/\nu$ (c is the velocity of light), and the wave vector $\mathbf{k}$, which determines the direction of the wave's propagation (the absolute value of $\mathbf{k}$ is $k = 2\pi/\lambda = = 2\pi\nu/c = \omega/c$). If we define $\hbar$ as *Planck's constant h* divided by 2π, that is, $h = 2\pi\hbar$, the energy of the photon, E, and its

momentum, **p**, will be related to the wave parameters as

$$E = \hbar\omega, \qquad \mathbf{p} = \hbar\mathbf{k} \quad (p = 2\pi\hbar/\lambda) \qquad (3.1)$$

where

$$\hbar = 1.054 \times 10^{-27}\,\mathrm{erg\,s} \qquad (3.2)$$

It follows from Eqs. (3.1) that using light of short wavelengths favourable for localizing a particle in ordinary space means using high-energy photons that are capable of transferring a great impact (momentum) to the particle and thereby upsetting its localization in momentum space. Using low-energy photons means using light of long wavelengths, and this in turn broadens the diffraction bands and reduces the precision of localizing a particle in ordinary (coordinate) space.

Equations (3.1) relate the wave properties of a photon to its corpuscular properties. Their right-hand members contain ω and **k**, which are determined by the diffraction pattern, and their left-hand members, E and **p**, describe the photon as a particle. Hence Eqs. (3.1) reflect the wave-corpuscular duality of a photon.

The *wave-corpuscular duality* proves to be a general property not only of photons but of all particles. This makes it possible to correlate the concepts of the electron as a particle and as a wave. The first to suggest the idea of the wave property of matter was Louis de Broglie, and proof came later when the diffraction of electrons was discovered. A more precise statement of this idea is contained in a proper interpretation of the mathematical apparatus of quantum mechanics.

We can express the results of Heisenberg's reasoning, just elaborated, concerning the limits of precision of measurement in the form of the following inequalities:

$$\Delta x\,\Delta p_x \geqslant \hbar, \qquad \Delta y\,\Delta p_y \geqslant \hbar, \qquad \Delta z\,\Delta p_z \geqslant \hbar \qquad (3.3)$$

in which Δx, Δy, Δz characterize the size of the region in coordinate space (x, y, z) containing the particle, and Δp_x, Δp_y, Δp_z the size of the region in momentum space (p_x, p_y, p_z) containing the particle. The inequalities show that the very nature of a particle makes it impossible to localize it simultaneously in coordinate space and in momentum space. They are called the *Heisenberg uncertainty relations*, or simply the *uncertainty relations*. The word "uncertainty" is understood to mean the regions of localization, $(\Delta x, \Delta y, \Delta z)$ and $(\Delta p_x, \Delta p_y, \Delta p_z)$, in the corresponding coordinate and momentum spaces.

We can couple the uncertainty relations (3.3) with

$$\Delta t\,\Delta(E' - E) \geqslant \hbar \qquad (3.4)$$

which links the uncertainty in the change of energy of a particle, $E' - E$, with the uncertainty in the time during which this change

occurred. According to (3.4), the transfer of energy cannot be localized precisely in time. Relation (3.4) can be called the *Heisenberg-Bohr uncertainty relation.*

The uncertainty relations (3.3) and (3.4) characterize the range of application of the classical (absolute) manner of describing phenomena. Since Planck's constant is small, this manner of description can unquestionably be used in referring to macroscopic bodies and their interactions. But this does not exhaust its significance. It is important in describing quantum processes because it is applied to the instruments used to study atomic objects. Experiments (with atomic objects as well) are always described in the classical (absolute) way.

The instruments and other means of observation, including the human senses (which, so to say, play the role of instruments built into the human body), are the necessary intermediaries between the human brain and the atomic object under consideration. We can now define more accurately what is meant by means of observation by indicating the manner of their description:

The means of observation must be described in classical terms but with due regard to the uncertainty relations (3.3) and (3.4).

4. Relativity with respect to the means of observation as the basis for the quantum way of describing phenomena

The new, quantum manner of describing phenomena must allow for the possibility of actual measurement of the properties of a micro-object. We must not ascribe to any object properties and states of motion that cannot be justified. For this reason particular attention should be given to the way in which we specify properties and states of motion. We must bear in mind the design and operation of the instruments that create the conditions to which the object is subjected. As has been said, the instruments and the external conditions must be described in the classical manner by indicating their parameters. It stands to reason that these parameters can be defined only to an accuracy permitted by the uncertainty relations. Otherwise we will be exceeding the actual potential of the measuring instruments.

A micro-object is revealed in its interaction with an instrument. For instance, the path of a charged particle becomes visible in the irreversible snowballing process that takes place in a cloud chamber or in the emulsion of a photographic plate (the particle loses its energy in ionizing the vapour or the chemicals of the emulsion; hence, its momentum becomes uncertain). The results of the interaction of an atomic object with a measuring instrument

(which is described classically) are the main experimental elements the systematization of which, based on the assumptions about the properties of the object, makes up the aim of the theory: from a study of such interactions we can deduce the properties of the atomic object, and the predictions of the theory are formulated as the expected results of these interactions.

Such a statement of the problem allows the introduction of quantities that characterize the object irrespective of the measuring instrument (electric charge, mass, and properties described by quantum mechanical operators) and at the same time makes possible a comprehensive approach to the object: the object can be viewed from the aspect (wave or corpuscular, for instance) necessitated by the instrument and by the external conditions the instrument creates.

The new statement of the problem makes it possible to consider the case when the various aspects and properties of an object do not manifest themselves simultaneously, that is, when *particularization of the object's behaviour* is impossible. This will be so if incompatible external conditions are needed for the manifestation of the object's properties (for instance, wave and corpuscular).

We can act on the proposal of Niels Bohr and call *complementary* the properties that reveal themselves in their pure form only in different experiments held in mutually exclusive conditions, whereas in conditions of one and the same experiment they manifest themselves only in an incomplete, modified form (for instance, the incomplete localization in the coordinate and the momentum space permitted by the uncertainty relations). There is no sense in considering complementary properties simultaneously (in the pure form), which explains the absence of a contradiction in the concept of wave-corpuscular duality.

By making the results of the interaction of a micro-object and a measuring instrument the basis of the new manner of description we introduced an important concept, the concept of *relativity with respect to the means of observation*, which generalizes the well-known concept of relativity with respect to the frame of reference. Such a manner of description does not at all mean that we are ascribing a lesser degree of reality to the micro-object than to the measuring instrument or that we are reducing the properties of the micro-object to the properties of the instrument. On the contrary, a description on the basis of the concept of relativity with respect to the means of observation gives a much deeper, more refined, and more objective picture of the micro-object than was possible on the basis of the idealizations of classical physics. Such a picture also requires a more sophisticated mathematical apparatus, namely, the theory of linear operators, including eigenfunctions and eigenvalues, the theory of groups, and other math-

ematical concepts. The use of this apparatus in quantum physics made it possible to give a theoretical explanation of some fundamental properties of matter that could not be explained in the classical way and also to calculate the values of many quantities observed in experiments (for instance, the frequencies in atomic spectra). But more than that — and this is no less important to us — the physical interpretation of the mathematical concepts used in quantum mechanics leads to a number of profound and principled conclusions; for one, generalization of the concept of the state of a system on the basis of the concepts of probability and potential possibility.

5. Potential possibility in quantum mechanics

If we take the act of interaction between an atomic object and a measuring instrument as the source of our judgements about the object's properties and if in studying phenomena we allow for the concept of relativity with respect to the means of observation, we are introducing a substantially new element into the description of the atomic object and its state and behaviour, that is, the idea of probability and thereby the idea of potential possibility. The need to consider the concept of probability as a substantial element of description rather than a sign of incompleteness of our knowledge follows from the fact that for given external conditions the result of the object's interaction with the instrument is not, generally speaking, predetermined uniquely but only has a certain probability of occurring. With a fixed initial state of the object and with given external conditions a series of such interactions results in a statistics that corresponds to a certain probability distribution. This probability distribution reflects the potential possibilities that exist in the given conditions.

Let us consider an experiment with a physical system that would enable us to make predictions about the results of future interactions between the system and measuring instruments of various kinds. Such an *initial experiment* must include a certain *preparation of the system* (for instance, preparation of a monochromatic beam of electrons) and the creation of certain external conditions in which the system will be placed after the preparation (for instance, the passage of the electron beam through a crystal). At times it is advisable to consider the preparation of the system and the creation of external conditions as two different stages of the experiment, but the two stages can also be considered one initial experiment, the purpose of which is to obtain predictions:

The initial experiment is always addressed to the future.

The manner of preparation and the external conditions in an initial experiment are described in the language of classical physics, but its result, which must give a full catalogue of the potential possibilities for the given conditions, requires new, quantum mechanical means for formulation. To have an idea of why we must use these means, let us consider how the potential possibilities existing in the given conditions materialize.

First of all we must bear in mind that a *final experiment*, in which the potential possibilities materialize, may be conducted in different ways: the registering instrument may be of different construction (as a rule, one excludes another). As in the initial experiment, the construction and operation of the instrument are described in the classical way. The different versions of the final experiment and the corresponding instruments can be characterized by the type of the quantity they measure (position, momentum, etc.).

Thus, with the initial experiment given, there is first of all a possibility of choosing different types of instruments for the final experiment. In any case,

The final experiment is always addressed to the past

(and not to the future in contrast to the initial experiment). It can be called the *verifying experiment* because it enables us to verify the predictions of the initial experiment.

Let us assume that the type of verifying experiment has been chosen. How do we formulate its result? We must always remember that we are talking about potential possibilities, which are created in the initial experiment and realized in the verifying experiment. For a given type of verifying experiment these potential possibilities are expressed as probability distributions for the given quantity (more precisely, for the values of the quantity that can be obtained in the verifying experiment). Hence it is the probability distribution we seek to verify. Clearly, this cannot be done by a simple measurement but requires many repetitions of the entire experiment (with the same preparation of the object and the same external conditions). The statistics obtained in this process of repetition makes it possible to draw a conclusion about the probability distribution that is to be studied.

A *total experiment* (an experiment that is carried out to the end and permits a comparison with theory) consists of the initial and verifying experiments combined and performed many times over. Here it is in place to note once more that for a given initial experiment (for given initial conditions) the final experiment may be set up in different ways (the measured quantities may differ)

and every type of final experiment has its own probability distribution.

Thus a theory must describe the initial state of a system in such a way as to make it possible to obtain probability distributions for any type of final experiment from this state. In this way we secure a full description of the potential possibilities that follow from the initial experiment.

Since a final experiment may take place later than the initial experiment, a theory must also give the time dependence of probabilities and potential possibilities. The establishment of this dependence will play the same role as the discovery of the laws of motion did in classical physics.

Chapter II

THE MATHEMATICAL APPARATUS
OF QUANTUM MECHANICS

1. Quantum mechanics
and the linear-operator problems

An important step towards the creation of present-day quantum mechanics was Bohr's postulation of two principles characterizing the properties of atomic systems.

The first principle asserts that atomic systems have stationary states, in which they do not radiate or absorb energy. In these states an atomic system possesses energy values that form a discrete sequence E_1, E_2, ..., E_n, ... (the *energy levels* of the system).

According to the second principle, radiation emitted or absorbed by an atomic system in the transition from one energy level to another has a definite frequency ν determined by the condition

$$E_m - E_n = h\nu$$

where h is Planck's constant, and E_m and E_n are the energy levels.

These principles conflict with classical mechanics and electrodynamics but are fully confirmed in experiments. It is a natural idea, therefore, to propose replacing the classical theory by a theory that would harmonize with Bohr's principles and be logically consistent.

The problem of determining the stationary states of an atomic system, states that are described by definite energy values (and certain other constants of integration), is analogous to the problem of mathematical physics where definite states of a system are chosen from the whole set of states, namely, the problem of eigenfrequencies of oscillations, or, more generally, the *linear-operator problem* and the associated *eigenvalue problem*. In the problem of this kind a sequence of values of a given quantity would emerge automatically from the whole set of values. Quantum mechanics has substantiated this idea of *quantization* ever since the historic paper of Erwin Schrödinger (1926) concerning quantization as an eigenvalue problem. A certain linear operator is related to each physical quantity, and the theory of linear operators is the mathematical apparatus of quantum mechanics.

2. The operator concept and examples

As in the case of a function, which is an instruction as to how, knowing number x, we can find number $y = f(x)$, an operator will map a given function $\varphi(x)$ into a new function

$$\psi(x) = L[\varphi(x)] \tag{2.1}$$

A *linear operator* has the properties that, for any functions φ_1, φ_2, φ,

$$L(\varphi_1 + \varphi_2) = L(\varphi_1) + L(\varphi_2)$$
$$L(a\varphi) = aL(\varphi) \tag{2.2}$$

where a is an arbitrary complex number. Since we will deal only with linear operators, the word "linear" will be often omitted.

Operators act on functions of one or several variables. The variables can be either continuous, which is the case for the coordinates (position) of an object, or discontinuous, that is, assuming only discrete values, which is the case for energy levels or the number that labels these levels. Continuous variables can either take on any value or change within certain domains. Discontinuous variables can take on both finite and infinite sequences of values. We will always assume that the values of the independent variables (or arguments) of a function are real numbers, whereas the functions themselves, which the operators act on, can be complex-valued. When specifying an operator, we must always indicate on the functions of what variables it acts.

Typical operators that act on functions of a continuous variable x are the multiplication of a function into x and the differentiation with respect to x:

$$L[f(x)] = xf(x), \qquad L[f(x)] = \frac{d}{dx} f(x)$$

In the first case x plays a double role: it is the argument of $f(x)$ and it is the operator itself.

Another example is the Laplacian operator ∇^2.

$$\nabla^2 f(x, y, z) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}$$

There is also a class of linear operators that can be represented by a definite integral:

$$L[f(x)] = \int_a^b K(x, \xi) f(\xi) \, d\xi \tag{2.3}$$

where the function $K(x, \xi)$ is called the *kernel* of the operator. As one example of a kernel let us consider Poisson's differential equation

$$\nabla^2 F = f$$

If $f(x, y, z)$ is specified in an unbounded region and if the asymptotic behaviour of the solution at infinity is $F = 0$, the solution of Poisson's equation is

$$F(x, y, z) = G[f(x, y, z)] = -\frac{1}{4\pi} \iiint \frac{f(\xi, \eta, \zeta)}{r} \, d\xi \, d\eta \, d\zeta$$

The operator G has a kernel

$$-\frac{1}{4\pi r} = -\frac{1}{4\pi \left[(x-\xi)^2 + (y-\eta)^2 + (z-\zeta)^2\right]^{1/2}}$$

If the equation

$$L(F) = f \tag{2.4}$$

and the appropriate asymptotic conditions yield

$$L^{-1}(f) = F \tag{2.5}$$

then it is said that the operator L has an *inverse* (*inverse operator*) L^{-1}. In our example G is the inverse of ∇^2.

If a variable takes on only discrete values, these values can always be labelled by positive integers. Hence, a function of a discontinuous variable can always be replaced by a function of a positive integer, the number labelling the value. Any operator that acts on the function f_n of a positive integer n (more exactly, the result of this action) can be represented in the form of a sum

$$Lf_n = \sum_m K_{nm} f_m \tag{2.6}$$

The totality of the K_{nm}'s is called the *matrix* of the operator, and the operator is said to be in the *matrix representation*. Formula (2.6) is in perfect agreement with formula (2.3), and the matrix (K_{nm}) is the kernel of this operator.

3. Hermitian conjugate. Hermiticity

To every linear operator L we can relate another operator L^+ such that satisfies a certain functional equation and is called the *hermitian conjugate* of L. [1] We will denote it by the same letter as the original operator but with a dagger as a superscript (L^+ is pronounced el-dagger).

We can define a hermitian conjugate operator as follows. Given two functions f and g that satisfy some general conditions but in all other respects are arbitrary. The functional equation that determines a hermitian conjugate operator L^+ is

$$\int \left[\bar{g}L(f) - \overline{L^+(g)}f\right] d\tau = 0 \tag{3.1}$$

[1] In mathematics the terms adjoint, conjugate, and associate operator are used.

if all independent variables are continuous. It is

$$\sum_{mn} \left[\bar{g}_n L\left(f_m\right) - \overline{L^+ \left(g_n\right) f_m} \right] = 0 \qquad (3.1^*)$$

if they are discontinuous. The integration in (3.1) and the summation in (3.1*) are over all values of the independent variables in a given range. The symbol $d\tau$ in (3.1) denotes the volume element in this range. By a bar we denote the complex conjugate. Now let us turn to the general conditions that f and g must satisfy. First, the sums and integrals in which f and g appear must be convergent. Second, f and g must satisfy certain boundary conditions, which, generally speaking, depend on the type of L.

If L^+ coincides with L, then L is called a *hermitian operator*.

When the independent variable is discontinuous, we can write L in the form (2.6), and formula (3.1*) then yields

$$\sum_{mn} \left(K_{nm} - \bar{K}^+_{mn} \right) \bar{g}_n f_m = 0 \qquad (3.2)$$

For this equation to hold for arbitrary f and g, the coefficient of every product $\bar{g}_n f_m$ must vanish. If we equate the complex conjugate of (3.2) with zero, we find the matrix elements of a hermitian conjugate operator:

$$K^+_{mn} = \bar{K}_{nm} \qquad (3.3)$$

An operator is hermitian if and only if its matrix elements satisfy the condition

$$K_{mn} = \bar{K}_{nm} \qquad (3.3^*)$$

Such a matrix is called *hermitian (self-adjoint, self-conjugate)*. If we arrange its elements in a rectangular array in such a way that, say, K_{ij} stands in the ith row and in the jth column, the elements in the principal diagonal (that is, with the same subscripts) are real, and any two matrix elements that are in a position symmetric with respect to the principal diagonal are complex conjugate quantities.

Let us now consider the case of one continuous variable and assume that an operator L has a kernel $K(x, \xi)$. We denote the kernel of the hermitian conjugate operator by $K^+(x, \xi)$ and use formula (2.3) to obtain condition (3.1) in the form

$$\iint \left[K\left(x, \xi\right) - \overline{K^+ \left(\xi, x\right)} \right] \bar{g}\left(x\right) f\left(\xi\right) dx \, d\xi = 0 \qquad (3.4)$$

which holds true for arbitrary functions f and g if and only if the expression in brackets vanishes. This yields the following expression for the kernel of the hermitian conjugate in terms of the

kernel of the original operator:

$$K^+(x, \xi) = \overline{K(\xi, x)} \tag{3.5}$$

Thus, knowing the kernel of an operator, we can find the kernel of the hermitian conjugate by interchanging the independent variables of the original kernel and then forming the complex conjugate. The condition for hermiticity is

$$K(x, \xi) = \overline{K(\xi, x)} \tag{3.5*}$$

In a similar manner we can find the kernel of a hermitian conjugate when there are several independent variables. Let us illustrate this with the help of the operator G of the previous section. The kernel of G, where G is the inverse of the Laplacian operator ∇^2 is invariant under the interchange of x, y, z and ξ, η, ζ. Since the kernel is also a real function, G is hermitian.

If the inverse operator is hermitian, the original is hermitian too. Notably, the Laplacian operator is hermitian. This can be proved directly using the general formula (3.1). In fact, using vector notation, we find that

$$\bar{g} \cdot \nabla^2 f - \nabla^2 \bar{g} \cdot f = \text{div} \, [\bar{g} \, \text{grad} \, f - (\text{grad} \, \bar{g}) f] \tag{3.6}$$

If, in addition, f and g vanish at infinity, then

$$\int [\bar{g} \cdot \nabla^2 f - \nabla^2 \bar{g} \cdot f] \, d\tau = \int \text{div} \, [\bar{g} \, \text{grad} \, f - (\text{grad} \, \bar{g}) f] \, d\tau = 0 \tag{3.7}$$

according to Gauss's integral theorem.

Let us consider another example. We set

$$Lf = \frac{\partial f}{\partial x} \tag{3.8}$$

and go on to find the hermitian conjugate of L. According to (3.1),

$$\int_a^b \left[\bar{g} \, \frac{\partial f}{\partial x} - \left(- \frac{\partial \bar{g}}{\partial x}\right) f\right] dx = \int_a^b \frac{\partial}{\partial x} (\bar{g} f) \, dx = 0 \tag{3.9}$$

if f and g vanish at the limits of integration. This yields

$$\overline{L^+ g} = - \frac{\partial \bar{g}}{\partial x}, \qquad L^+ f = - \frac{\partial f}{\partial x} \tag{3.10}$$

We see that L is not hermitian. But if we multiply it by a pure imaginary number, say $-i$, the new operator L_1 yields

$$L_1 f = - i \frac{\partial f}{\partial x} \tag{3.11}$$

which means that L_1 is hermitian.

4. Operator and matrix multiplication

The *product* of two operators K and L is an operator that consists in applying K and L successively. If we apply L first and K second, the product is written as

$$M = KL$$

But if we apply K first and L second, the product will be

$$N = LK$$

Generally speaking, M and N are different operators, which implies that the product of operators depends on the order of multiplication. For instance, if K denotes multiplication by x and L differentiation with respect to x,

$$Kf = x \cdot f, \qquad Lf = \frac{\partial f}{\partial x} \tag{4.1}$$

we find that the product M acts on f as

$$Mf = KLf = x\frac{\partial f}{\partial x} \tag{4.2}$$

whereas

$$Nf = LKf = \frac{\partial}{\partial x}(xf) = x\frac{\partial f}{\partial x} + f \tag{4.3}$$

We see that in our example $KL \neq LK$ and that

$$(LK - KL)f = \frac{\partial}{\partial x}(xf) - x\frac{\partial f}{\partial x} = f$$

Hence the difference $LK - KL$ is an *identity operator:*

$$LK - KL = 1 \tag{4.4}$$

In some cases, however, the product does not depend on the order of multiplication. Then operator multiplication becomes *commutative* and the operators are said to *commute*. For example, differentiation with respect to two different independent variables of a function is commutative, and the operators corresponding to the differentiations will commute.

Let us find the hermitian conjugate of

$$M = KL \tag{4.5}$$

We have

$$\int \bar{g} Mf \, d\tau = \int \bar{g} K(Lf) \, d\tau$$

Let us denote

$$Lf = f'$$

The previous expression will then be

$$\int \bar{g} K f' \, d\tau = \int \overline{(K^+ g)} f' \, d\tau$$

according to the definition of K^+. We then assume that

$$K^+ g = g'$$

which yields

$$\int \overline{(K^+ g)}\, f'\, d\tau = \int \bar{g}' L f\, d\tau = \int \overline{(L^+ g')}\, f\, d\tau$$

according to the definition of L^+. Substituting $K^+ g$ for g', we finally obtain

$$\int \bar{g} K L f\, d\tau = \int \overline{(L^+ K^+ g)}\, f\, d\tau \qquad (4.6)$$

If we compare (4.6) with the definition of a hermitian conjugate operator

$$\int \bar{g} M f\, d\tau = \int \overline{(M^+ g)}\, f\, d\tau$$

we find that

$$M^+ = L^+ K^+ \qquad (4.7)$$

or

$$(KL)^+ = L^+ K^+ \qquad (4.8)$$

Thus a hermitian conjugate of a product of operators is equal to the product of the hermitian conjugates in a reversed order.

If K and L are hermitian, their products are *not*, generally speaking, hermitian because

$$(KL)^+ = LK \neq KL$$

But if K and L are hermitian *and* commute, their products are hermitian.

Let us consider the product of two operators with kernels $K(x, \xi)$ and $L(x, \xi)$. We have

$$L f = \int L(x, \xi) f(\xi)\, d\xi, \qquad KL f = \int \left[\int K(x, \xi_1) L(\xi_1, \xi)\, d\xi_1 \right] f(\xi)\, d\xi$$

We first integrate with respect to ξ_1 and then, introducing the notation

$$KL(x, \xi) = \int K(x, \xi_1) L(\xi_1, \xi)\, d\xi_1 \qquad (4.9)$$

we find that the previous formula can be written as

$$KL f = \int KL(x, \xi) f(\xi)\, d\xi \qquad (4.10)$$

Thus the product $\underline{KL}$ has a kernel defined by (4.9).

We now go on to the case when the two operators in a product act on a function of a discontinuous variable. As we know, the operators can then be represented by matrices. We have

$$(Lf)_n = \sum_i L_{ni}f_i, \qquad (Kg)_n = \sum_j K_{nj}g_j \qquad (4.11)$$

If we assume that

$$g = Lf \qquad (4.12)$$

then

$$(KLf)_n = \sum_j \sum_i K_{nj}L_{ji}f_i$$

or

$$(KLF)_n = \sum_i (KL)_{ni}\,f_i \qquad (4.13)$$

where

$$(KL)_{ni} = \sum_j K_{nj}L_{ji} \qquad (4.14)$$

Equation (4.14) defines *matrix multiplication.* Thus $(KL)_{ni}$ is the *inner product* of the nth row of K and the ith column of L.

Let us consider a matrix U whose elements U_{ij} satisfy the conditions

$$\sum_j U^+_{ij}U_{jk} = \delta_{ik}, \qquad \sum_j U_{ij}U^+_{jk} = \delta_{ik} \qquad (4.15)$$

where, according to the notation in (3.3),

$$U^+_{ij} = \bar{U}_{ji} \qquad (4.16)$$

and the Kronecker delta, δ_{ik}, is 1 when $i = k$ and zero otherwise. If we use matrix multiplication, we can write (4.15) as

$$U^+U = 1, \qquad UU^+ = 1 \qquad (4.17)$$

where the "1" stands for the *identity matrix.* A matrix that satisfies Eqs. (4.17) is called *unitary,* and the corresponding operator is said to be unitary. A unitary operator has the following property. If

$$g = Uf \qquad (4.18)$$

or in matrix form

$$g_n = \sum_i U_{ni}f_i \qquad (4.18^*)$$

then

$$\sum_n \bar{g}_n g_n = \sum_{ijn} U^+_{in}U_{nj}\bar{f}_i f_j$$

But according to (4.15), we have

$$\sum_n \bar{g}_n g_n = \sum_i \bar{f}_i f_i \tag{4.19}$$

which means that a sum of type (4.19) is invariant under a unitary transformation.

5. Eigenvalues and eigenfunctions

The main problem in the theory of operators is the analysis of the equation

$$Lf = \lambda f \tag{5.1}$$

where λ is a constant quantity. The operator L is assumed to be a *normal* operator, that is, an operator that satisfies the condition

$$LL^+ = L^+L \tag{5.2}$$

Hermitian and unitary operators are normal.

A linear equation that is satisfied by assuming the unknown function to be zero is called a *homogeneous* linear equation, which is the case with Eq. (5.1). When analyzing such an equation, we must consider the *boundary conditions* as well. These conditions are also *homogeneous*, that is, such that $f = 0$ is a boundary condition. A homogeneous linear equation along with linear homogeneous boundary conditions comprise the *homogeneous problem*.

For arbitrary values of the parameter λ the homogeneous problem has, generally speaking, none but the trivial solution $f = 0$. Only when the parameter takes on very special values is a nontrivial solution possible. These values can form a denumerable or countably infinite set $\lambda_0, \lambda_1, \lambda_2, \ldots$ or can assume all values in a certain interval. The special values of λ are called *eigenvalues*, and the associated solutions of the homogeneous problem are called *eigenfunctions*. In mathematics one can come across *characteristic* values (functions) or *proper* values (functions), but not in physics. The complete set of eigenvalues constitutes the *spectrum* of the operator. A denumerable set is said to form a *discrete* spectrum, and when the eigenvalues assume all values in a certain interval, the spectrum is called *continuous*. In more complicated situations the spectrum may exhibit both discrete and continuous parts.

Eigenfunctions corresponding to a discrete spectrum have the property that

$$\sum \bar{f}f \quad \text{and} \quad \int \bar{f}f \, d\tau$$

over all values of the independent variables converge, whereas for the eigenfunctions of a continuous spectrum both the sum and the integral become infinite. In the latter case instead of considering the eigenfunctions themselves, we can integrate them with respect to λ over an infinitesimal interval of the continuous spectrum. If we then substitute these integrals for the functions in the above products, the sum and integral are finite.

For the eigenvalues of a hermitian operator we can prove the following theorem:

The eigenvalues of a hermitian operator are real.

Let f be a solution of (5.1). We multiply both sides of Eq. (5.1) by $\bar{f}$ and sum (or integrate) over all values of the independent variables and get

$$\sum \bar{f} L f = \lambda \sum \bar{f} f$$

or

$$\int \bar{f} L f \, d\tau = \lambda \int \bar{f} f \, d\tau$$

which yields

$$\lambda = \frac{\sum \bar{f} L f}{\sum \bar{f} f} \qquad (5.3)$$

or

$$\lambda = \frac{\int \bar{f} L f \, d\tau}{\int \bar{f} f \, d\tau} \qquad (5.3^*)$$

The denominators in the last two expressions are real positive numbers. We will next show that the numerators are also real. In fact, the imaginary parts of the numerators are

$$\frac{1}{2i} \sum [\bar{f} L f - \overline{(Lf)} f] \quad \text{or} \quad \frac{1}{2i} \int [\bar{f} L f - \overline{(Lf)} f] \, d\tau$$

These two vanish by (3.1). Thus the numerators in (5.3) and (5.3*) are real, which means that λ is real as well. The proof is complete.

In some cases formulas (5.3) and (5.3*) indicate the sign of the eigenvalues. For the Laplacian operator, (5.3*) yields

$$\lambda = \frac{\int \bar{f} \nabla^2 f \, d\tau}{\int \bar{f} f \, d\tau} = -\frac{\int (\text{grad } \bar{f} \cdot \text{grad } f) \, d\tau}{\int \bar{f} f \, d\tau}$$

Both integrals in the right-hand side are positive. Hence, the minus before the fraction indicates that the eigenvalues of the Laplacian operator are negative.

There will be many problems in finding eigenvalues and eigenfunctions throughout this book, so that there is no need to examine them further at this point.

6. The Stieltjes integral and the operator corresponding to multiplication into the independent variable

Let us first recall the concept of the *definite (Riemann) integral.* Suppose that the closed interval defined by the limits of integration λ_0 and λ is divided into n subintervals by the numbers $\lambda_1, \lambda_2, \ldots, \lambda_{n-1}$. (Here $\lambda_n = \lambda_1$) If we then increase the number of these subintervals indefinitely $(n \to \infty)$ so that even the largest tends to zero, the Riemann integral of $f(\lambda)$ over the closed interval $[\lambda_0, \lambda]$ is defined as

$$\int_{\lambda_0}^{\lambda} f(\lambda) \, d\lambda = \lim_{n \to \infty} \sum_{i=1}^{n} f(\lambda_i) \Delta\lambda_i \tag{6.1}$$

where $\Delta\lambda_i = \lambda_i - \lambda_{i-1}$.

We now go on to the *Stieltjes integral.* Let $\rho(\lambda)$ be either a monotonically increasing function or the difference between two monotonic functions. We construct the sum

$$\sum_{i=1}^{n} f(\lambda_i) \Delta\rho(\lambda_i)$$

where $\Delta\rho(\lambda_i) = \rho(\lambda_i) - \rho(\lambda_{i-1})$. When the process of subdivision of $[\lambda_0, \lambda]$ is indefinite, the Stieltjes integral is the limit of this sum:

$$\int_{\lambda_0}^{\lambda} f(\lambda) \, d\rho(\lambda) = \lim_{n \to \infty} \sum_{i=1}^{n} f(\lambda_i) \Delta\rho(\lambda_i) \tag{6.2}$$

Whenever $\rho(\lambda)$ is continuous and has a bounded first derivative, we can set

$$\Delta\rho(\lambda_i) = \rho'(\lambda_i) \Delta\lambda_i$$

which holds true since terms in $(\Delta\lambda_i)^2$ and higher powers vanish when passing to the limit. Then

$$\int_{\lambda_0}^{\lambda} f(\lambda) \, d\rho(\lambda) = \int_{\lambda_0}^{\lambda} f(\lambda) \rho'(\lambda) \, d\lambda \tag{6.3}$$

and (6.2) includes the Riemann integral (6.1) as a special case. But the Stieltjes integral has a meaning even when $\rho(\lambda)$ is a discontinuous function, that is, possesses finite discontinuities.

This is the case with the operator corresponding to multiplication into the independent variable if the variable is continuous.

Let us take a close look at such an operator. The corresponding eigenvalue equation is

$$xf(x, \lambda) = \lambda f(x, \lambda) \tag{6.4}$$

Now this equation cannot be satisfied by any "normal" type of function, since we would have to assume that $f(x, \lambda)$ is zero for all values of x except $x = \lambda$. Because of this let us formally integrale (6.4) with respect to λ:

$$x \int_{\lambda_0}^{\lambda} f(x, \lambda) d\lambda = \int_{\lambda_0}^{\lambda} \lambda f(x, \lambda) d\lambda \tag{6.5}$$

We denote

$$\int_{\lambda_0}^{\lambda} f(x, \lambda) d\lambda = F(x, \lambda) \tag{6.6}$$

and write (6.5) as

$$x \int_{\lambda_0}^{\lambda} d_\lambda F(x, \lambda) = \int_{\lambda_0}^{\lambda} \lambda \, d_\lambda F(x, \lambda) \tag{6.7}$$

where the integral on the right is the Stieltjes integral. Equation (6.7) is solved by a function

$$F(x, \lambda) = 1, \qquad \lambda \geqslant x$$
$$F(x, \lambda) = 0, \qquad \lambda < x \tag{6.8}$$

Indeed, all increments

$$\Delta F(x, \lambda_i) = F(x, \lambda_{i+1}) - F(x, \lambda_i)$$

except one are zero. The increment that is not zero is equal to unity and corresponds to the values of λ_i and λ_{i+1} that satisfy the inequality

$$\lambda_i < x < \lambda_{i+1}$$

When $[\lambda_0, \lambda]$ is subdivided indefinitely, this particular λ_i tends to x, and in the limit Eq. (6.7) holds true.

Thus our operator has no eigenfunctions in the conventional sense, but a function $F(x, \lambda)$ defined by (6.6) does exist.

If the independent variable x is discontinuous ar d assumes a sequence of discrete values

$$x_1, \ x_2, \ \ldots, \ x_n, \ \ldots$$

then the eigenvalue equation (6.4) must be written in the form of a set of equations

$$x_n f(x_n, \lambda) = \lambda f(x_n, \lambda), \qquad n = 1, 2, \ldots \qquad (6.9)$$

Equations (6.9) have a solution in the ordinary sense. Let

$$f(x_n, \lambda) = 1, \qquad \lambda = x_n$$
$$f(x_n, \lambda) = 0, \qquad \lambda \neq x_n \qquad (6.10)$$

This is obviously a solution to (6.9). The eigenvalues are

$$\lambda_n = x_n \qquad (6.11)$$

and the corresponding eigenfunctions are

$$f(x_n, \lambda_m) = \delta_{nm} \qquad (6.10^*)$$

7. Orthogonality of eigenfunctions and normalization

Let us examine a hermitian operator with a discrete spectrum. We write down the eigenvalue equations for f_n and f_m corresponding to two different eigenvalues λ_n and λ_m:

$$Lf_m = \lambda_m f_m, \qquad Lf_n = \lambda_n f_n \qquad (7.1)$$

For hermitian operators we have

$$\int [\bar{f}_n Lf_m - \overline{(Lf_n)} f_m] \, d\tau = 0 \qquad (7.2)$$

If we use (7.1) in (7.2), we obtain

$$(\lambda_m - \lambda_n) \int \bar{f}_n f_m \, d\tau = 0 \qquad (7.3)$$

Since by definition $\lambda_m \neq \lambda_n$,

$$\int \bar{f}_n f_m \, d\tau = 0, \qquad n \neq m \qquad (7.4)$$

This property of eigenfunctions is called *orthogonality*. Thus

Eigenfunctions corresponding to different eigenvalues of a hermitian operator are mutually orthogonal.

Since an eigenfunction satisfies a homogeneous equation, if it is multiplied by a factor, the product will also satisfy this equation. The factor can be chosen so that

$$\int \bar{f}_n f_n \, d\tau = 1 \qquad (7.5)$$

Making this choice is called *normalization*, and functions that satisfy (7.5) are spoken of as *normalized*. The normalization is not exactly definite because if we substitute $f'_n = e^{i\alpha_n} f_n$ for f_n,

with α_n real, $\bar{f}_n$ will be replaced by $\bar{f}'_n = e^{-i\alpha_n}\bar{f}_n$ and condition (7.5) will again hold.

The orthogonality condition and the normalization of functions can be written compactly as

$$\int \bar{f}_n f_m \, d\tau = \delta_{nm} \tag{7.6}$$

An eigenvalue can relate to one or several eigenfunctions, in fact, to as many as the equation

$$Lf = \lambda_n f$$

has linearly independent solutions for the given eigenvalue λ_n. When there are several eigenfunctions corresponding to one eigenvalue, we have *degeneracy*.

The number of solutions (denoted by s) may depend on n. If the solutions are

$$f_{n1}, \, f_{n2}, \, \ldots, \, f_{ns} \tag{7.7}$$

then any superposition

$$f_n = a_1 f_{n1} + a_2 f_{n2} + \ldots + a_s f_{ns} \tag{7.8}$$

will also be a solution. The functions (7.7) are not necessarily orthogonal. But we can always replace them with superpositions of type (7.8) that are orthogonal and normalized. Let us assume that this has been done. Then f_{nl} and f_{nk} will for $n = m$ satisfy the condition

$$\int \bar{f}_{nl} f_{nk} \, d\tau = \delta_{lk}, \qquad l, \, k = 1, 2, \ldots, s \tag{7.9}$$

At times it is convenient to denote all functions (7.7) by one symbol f_n. Then (7.9) can be written as (7.5), and its two sides are understood to be matrices with elements (7.9).

Condition (7.9) does not fully determine the function f_{nl}. In fact, if we set

$$f'_{nk} = \sum_{l=1}^{s} a_{kl} f_{nl} \tag{7.10}$$

with a_{kl} complying with the condition

$$\sum_{l=1}^{s} \bar{a}_{kl} a_{ll} = \delta_{kl} \tag{7.11}$$

then (7.9) will again hold. Matrix A with elements a_{ki} that satisfy (7.11) is said to be *unitary*. So is the transformation that is carried out by means of this matrix. We can thus say that the eigenfunctions corresponding to a degenerate eigenvalue can be determined up to a *unitary transformation*.

Let us now turn to the case of an operator with a continuous spectrum. The eigenvalue equation is

$$Lf(x, \lambda) = \lambda f(x, \lambda) \tag{7.12}$$

We integrate both sides with respect to λ twice: first over the interval $[\lambda_1, \lambda_1 + \Delta_1\lambda]$, then over the interval $[\lambda_2, \lambda_2 + \Delta_2\lambda]$. If we assume, as was done in considering the Stieltjes integral, that

$$F(x, \lambda) = \int_{\lambda_0}^{\lambda} f(x, \lambda) \, d\lambda \tag{7.13}$$

we find that

$$L[\Delta_1 F(x, \lambda)] = \int_{\lambda_1}^{\lambda_1 + \Delta_1\lambda} \lambda \, d_\lambda F(x, \lambda) \tag{7.14}$$

and

$$L[\Delta_2 F(x, \lambda)] = \int_{\lambda_2}^{\lambda_2 + \Delta_2\lambda} \lambda \, d_\lambda F(x, \lambda) \tag{7.14*}$$

where for the sake of brevity we have set

$$\Delta_k F(x, \lambda) = F(x, \lambda_k + \Delta_k\lambda) - F(x, \lambda_k), \quad k = 1, 2 \tag{7.15}$$

A quantity defined by (7.15) is called a *proper differential*. Proper differentials, with respect to x, are square-integrable, whereas the functions $f(x, \lambda)$ are not.

We multiply (7.14) by $\overline{\Delta_2 F}$ and the complex conjugate of (7.14*) by $\Delta_1 F$, subtract one from the other, and integrate the result. The left-hand side nullifies, which means that the right-hand side is also zero, that is,

$$\int d\tau \int_{\lambda_2}^{\lambda_2 + \Delta_2\lambda} \int_{\lambda_1}^{\lambda_1 + \Delta_1\lambda} (\lambda - \mu) \, d_\mu \overline{F(x, \mu)} \, d_\lambda F(x, \lambda) = 0 \tag{7.16}$$

The last equation is valid for any λ_1, λ_2, $\Delta_1\lambda$, and $\Delta_2\lambda$. Let us assume that the intervals $\Delta_1\lambda$ and $\Delta_2\lambda$ are separated by a finite distance and that both are infinitesimal quantities. This means that up to infinitesimals the difference $\lambda - \mu$ is equal to $\lambda_1 - \lambda_2$. Since the latter is nonzero, we can cancel it out and obtain

$$\int \overline{\Delta_2 F} \, \Delta_1 F \, d\tau = 0 \tag{7.17}$$

Thus

Proper differentials corresponding to different intervals are mutually orthogonal.

We now assume that $\Delta_1\lambda$ and $\Delta_2\lambda$ coincide, and we consider the integral

$$J = \int \overline{\Delta F}\, \Delta F\, d\tau \qquad (7.18)$$

We chose arbitrarily two eigenvalues λ_1 and λ_2 such that the interval $\Delta\lambda$ lies between them:

$$\lambda_1 < \lambda < \lambda + \Delta\lambda < \lambda_2$$

By the orthogonality condition for proper differentials the integral J will not change if we add to it

$$\int \overline{\Delta F}\left(\int_{\lambda_1}^{\lambda} f(x,\lambda)\,d\lambda\right)d\tau + \int \Delta F\left(\int_{\lambda+\Delta\lambda}^{\lambda_2} f(x,\lambda)\,d\lambda\right)d\tau$$

Hence,

$$J = \int \overline{\Delta F}\,[F(x,\lambda_2) - F(x,\lambda_1)]\,d\tau \qquad (7.19)$$

which implies that J is an infinitesimal of the first order relative to $\Delta\lambda$, rather than of the second order as one would expect it to be. The integral (7.18) can be normalized so that

$$\lim_{\Delta\lambda\to 0}\frac{1}{\Delta\lambda}\int |\Delta F|^2\, d\tau = 1 \qquad (7.20)$$

This is the usual normalization conditions for the "eigenfunctions" of an operator with a continuous spectrum. For instance, the function $F(x,\lambda)$ of Section 6 can be used for normalization.

8. Expansion in eigenfunctions.
Completeness property of eigenfunctions

We will consider the eigenfunctions

$$f_n(x) = f(x,\lambda_n) \qquad (8.1)$$

of an operator with a discrete spectrum and assume them to be normalized. Let $f(x)$ be a square-integrable arbitrary function. We will try to expand it in a series of functions $f_n(x)$. For this we set

$$f(x) = \sum_{k=0}^{n} a_k f_k(x) + R_n(x) \qquad (8.2)$$

The sum of the products on the right side constitutes the first n terms of the expansion, and $R_n(x)$ denotes the remainder terms after the nth term. The expansion coefficients a_k are chosen in such a way as to insure the smallest possible error, where for

the measure of error we take the integral

$$\rho_n = \int |R_n(x)|^2 \, dx \tag{8.3}$$

which is

$$\rho_n = \int \left| f(x) - \sum_{k=0}^{n} a_k f_k(x) \right|^2 dx \tag{8.4}$$

This integral is a quadratic function in the n unknowns a_k. We will look for the minimum of (8.4), and for this will equate with zero the first partial derivatives with respect to the expansion coefficients. By the orthogonality condition for eigenfunctions we find that

$$\rho_n = \int |f(x)|^2 \, dx - \sum_{k=0}^{n} a_k \int \overline{f(x)} \, f_k(x) \, dx$$

$$- \sum_{k=0}^{n} \bar{a}_k \int f(x) \, \overline{f_k(x)} \, dx + \sum_{k=0}^{n} \bar{a}_k a_k \tag{8.5}$$

Obviously we can differentiate with respect to $\bar{a}_k$ and a_k as if these were independent quantities. Equating the derivative with respect to $\bar{a}_k$ with zero, we obtain the following formula for a_k:

$$a_k = \int \overline{f_k(x)} \, f(x) \, dx \tag{8.6}$$

The use of (8.6) in the expression (8.5) for the mean square error ρ_n yields

$$\rho_n = \int |f(x)|^2 \, dx - \sum_{k=0}^{n} |a_k|^2 \tag{8.7}$$

Since by definition ρ_n cannot be negative, for any n there is the inequality

$$\sum_{k=0}^{n} |a_k|^2 \leqslant \int |f(x)|^2 \, dx \tag{8.8}$$

If for any square-integrable function $f(x)$

$$\lim_{n \to \infty} \rho_n = 0 \tag{8.9}$$

or, which is the same,

$$\sum_{k=0}^{\infty} |a_k|^2 = \int |f(x)|^2 \, dx \tag{8.10}$$

the set of functions $f_k(x)$ is said to be *complete*, which means that no function $f(x)$ can be found that would be orthogonal to all functions $f_k(x)$. Indeed, for such a function all expansion coefficients are zero, which implies that $\int |f(x)|^2 dx$ is also zero. But this is possible only if $f(x)$ is zero (with the possible exception of certain values of x).

Equation (8.9) shows that in the limit the remainder $R_n(x)$ vanishes (with the possible exception of certain values of x), so we have here a series expansion

$$f(x) = \sum_{n=0}^{\infty} a_n f_n(x) \tag{8.11}$$

If a second function $g(x)$ can be expanded in a series

$$g(x) = \sum_{n=0}^{\infty} b_n f_n(x) \tag{8.12}$$

then we have

$$\int \overline{f(x)} g(x) dx = \sum_{n=0}^{\infty} \bar{a}_n b_n \tag{8.13}$$

which is a generalization of (8.10).

There is a theorem in mathematics for operators of a very general form, which states that

The totality of eigenfunctions makes up a complete set of functions.

If there is a degeneracy, in the expansion

$$f(x) = \sum_{n=0}^{\infty} a_n f_n(x) \tag{8.11}$$

every term

$$a_n f_n(x) \tag{8.14}$$

must be replaced by a sum

$$a_{n1} f_{n1}(x) + a_{n2} f_{n2}(x) + \ldots + a_{ns} f_{ns}(x) \tag{8.14*}$$

where

$$a_{nl} = \int \overline{f_{nl}(x)} f(x) dx \tag{8.15}$$

In the case of a continuous spectrum the proper differentials divided by $(\Delta\lambda)^{1/2}$,

$$\frac{1}{(\Delta\lambda)^{1/2}} \Delta F(x, \lambda) \tag{8.16}$$

form a mutually orthogonal and normalized set of functions. Hence we can write the expansion in the form

$$f(x) = \sum_{\lambda} a(\lambda) \frac{1}{(\Delta\lambda)^{1/2}} \Delta F(x, \lambda) \tag{8.17}$$

where

$$a(\lambda) = \frac{1}{(\Delta\lambda)^{1/2}} \int \overline{\Delta F(x, \lambda)} f(x) \, dx \tag{8.18}$$

or, as $\Delta\lambda \to 0$,

$$f(x) = \int c(\lambda) \, d_{\lambda} F(x, \lambda) \tag{8.19}$$

where

$$c(\lambda) = \lim_{\Delta\lambda \to 0} \frac{1}{\Delta\lambda} \int \overline{\Delta F(x, \lambda)} f(x) \, dx \tag{8.20}$$

At times these formulas can be replaced by simpler ones:

$$f(x) = \int c(\lambda) f(x, \lambda) \, d\lambda, \qquad c(\lambda) = \int \overline{f(x, \lambda)} f(x) \, dx \tag{8.21}$$

These represent an integral expansion similar to Fourier integral expansion, which is a particular case of (8.21).

The completeness property of eigenfunctions for an operator with a continuous spectrum is

$$\int |f(x)|^2 \, dx = \int |c(\lambda)|^2 \, d\lambda \tag{8.22}$$

or, if $b(\lambda)$ is an expansion coefficient of another function $g(x)$ (just as $c(\lambda)$ is for $f(x)$),

$$\int \overline{f(x)} \, g(x) \, dx = \int \overline{c(\lambda)} \, b(\lambda) \, d\lambda \tag{8.23}$$

If an operator has a continuous spectrum as well as a discrete spectrum, the expansion in eigenfunctions and the completeness condition will each include both a sum and an integral.

Chapter III

QUANTUM MECHANICAL OPERATORS

1. Interpretation of the eigenvalues of an operator

At the beginning of the previous chapter we found that in quantum mechanics a definite linear operator is related to each physical quantity. What is the meaning of such a relationship? We saw that an operator possesses certain eigenvalues and eigenfunctions. Now we must clarify the physical meaning of these mathematical concepts. Let us start by interpreting the eigenvalues, they being a simpler concept. We will introduce the following hypothesis:

The eigenvalues of an operator related to a given physical quantity are the values that this quantity assumes in the conditions created by measuring it.

We must note the importance of specifying the conditions in which the quantity assumes its values. If we measure a quantity that is not included in the same group with the original quantity (see Chapter I), new conditions are created. In these the original quantity may not have definite values. However, when measuring a quantity, we create conditions in which one of the eigenvalues of the corresponding operator must appear as the result of the measuring process. To state our interpretation more briefly, we can say: The eigenvalues of an operator are the values of the corresponding physical quantity.

Hence there is a limitation on the form of the operator corresponding to a real physical quantity. Since all the values of this quantity are real, the operator must have only real eigenvalues, which points to its hermiticity. Thus

A real physical quantity is described by a hermitian operator.

As we know, an operator can have both discrete and continuous spectra. Therefore operators can correspond to quantities that assume both a denumerable set of values and a set of all the values in a certain interval. We must note here that the old quantum theory could formulate "quantum conditions" only for quantities that change abruptly and did not include the cases when quantities change continuously.

2. Poisson brackets

How do we find the operator for a given physical quantity? We may use two guiding principles. First, the eigenvalue spectrum of the operator must coincide with the spectrum of observable values of the physical quantity. Second, the relations between operators must correctly reflect relations between the quantities. In relating operators to physical quantities we must note that the analogy with classical mechanics plays an important role. But this *classical analogy* must be used cautiously since it may be incomplete.

In classical mechanics a system is described by canonical variables: the generalized coordinates $q_1, q_2, \ldots, q_n$ and the generalized momenta $p_1, p_2, \ldots, p_n$. We can define *canonically conjugate variables* with the help of the *Poisson bracket*. We start with *Hamilton's canonical equations of motion*

$$\frac{dq_k}{dt} = \frac{\partial H}{\partial p_k}, \quad \frac{dp_k}{dt} = -\frac{\partial H}{\partial q_k}, \quad k = 1, 2, \ldots, n \qquad (2.1)$$

Let F be a certain function of coordinates, momenta, and time:

$$F = F(q_1, q_2, \ldots, q_n, p_1, p_2, \ldots, p_n, t) \qquad (2.2)$$

We construct the total time-derivative of F

$$\frac{dF}{dt} = \frac{\partial F}{\partial t} + \sum_{k=1}^{n} \left(\frac{\partial F}{\partial q_k} \frac{dq_k}{dt} + \frac{\partial F}{\partial p_k} \frac{dp_k}{dt} \right) \qquad (2.3)$$

Using Eqs. (2.1) in (2.3), we obtain

$$\frac{dF}{dt} = \frac{\partial F}{\partial t} + [H, F] \qquad (2.4)$$

where we call

$$[H, F] = \sum_{k=1}^{n} \left(\frac{\partial H}{\partial p_k} \frac{\partial F}{\partial q_k} - \frac{\partial H}{\partial q_k} \frac{\partial F}{\partial p_k} \right) \qquad (2.5)$$

the *classical* Poisson bracket of the functions H and F. Likewise, for any pair of functions F and G the Poisson bracket is

$$[F, G] = \sum_{k=1}^{n} \left(\frac{\partial F}{\partial p_k} \frac{\partial G}{\partial q_k} - \frac{\partial F}{\partial q_k} \frac{\partial G}{\partial p_k} \right) \qquad (2.6)$$

The main property of the classical Poisson bracket is its invariance under a *contact transformation*, that is, a transformation of the variables p_k and q_k that leaves the form of Hamilton's equations unaltered. Furthermore, the Poisson bracket has the

following properties (which are easily derived from its definition):

$$[F, G] = - [G, F] \tag{2.7}$$

$$[F, c] = 0 \tag{2.8}$$

with c being a constant not depending on p_k or q_k.

Also

$$[F_1 + F_2, G] = [F_1, G] + [F_2, G] \tag{2.9}$$

$$[F_1 F_2, G] = F_1 [F_2, G] + [F_1, G] F_2 \tag{2.10}$$

Finally, we have the identity

$$[F, [G, L]] + [G, [L, F]] + [L, [F, G]] \equiv 0 \tag{2.11}$$

The Poisson brackets for generalized coordinates and momenta are

$$[q_k, q_l] = 0, \qquad [p_k, p_l] = 0, \qquad [p_k, q_l] = \delta_{kl} \tag{2.12}$$

In classical physics (2.12) can serve as a definition of canonically conjugate coordinates and momenta.

As we have already noted, all relationships that use the Poisson bracket (for instance, the formula for the total time-derivative of a function) do not depend on the choice of the generalized coordinates and momenta. We can thus expect that because there is an analogy between classical and quantum mechanics, there should be something similar to the Poisson bracket in quantum mechanics.

The form of the *quantum* Poisson bracket was found by P. A. M. Dirac on the basis of *Bohr's correspondence principle* with the classical formula (2.6) as the starting point. Our approach will be different. Essentially it belongs to Dirac and is based on the assumption that the quantum Poisson bracket of any two noncommutative operators possesses all the properties (2.7)-(2.11).

We start from Eq. (2.10), where we substitute G_1 and G_2 for F_1 and F_2, and F for G, and use Eq. (2.7). As a result we obtain another equation:

$$[F, G_1 G_2] = G_1 [F, G_2] + [F, G_1] G_2 \tag{2.10*}$$

We will think of F and G as noncommutative operators. Hence the order of multiplication in (2.10) is significant. Let us assume that the order of multiplication is that as in (2.10). This can be justified in the following manner. If $G = H$. Eq. (2.10) corresponds, at least in classical mechanics, to the rule of finding the time derivative of $F_1 F_2$. When we are dealing with noncommutative operators and wish to find the time derivative of $F_1 F_2$, we must keep the order of multiplication as in (2.10), with F_1 always to the left of F_2.

Now let us put $G = G_1 G_2$ in (2.10). Using (2.10*), we can write (2.10) in the form

$$[F_1 F_2,\ G_1 G_2] = F_1 G_1 [F_2,\ G_2] + F_1 [F_2,\ G_1] G_2$$
$$+ G_1 [F_1,\ G_2] F_2 + [F_1,\ G_1] G_2 F_2 \qquad (2.13)$$

On the other hand, setting $F = F_1 F_2$ in (2.10*) and using (2.10), we obtain

$$[F_1 F_2,\ G_1 G_2] = G_1 F_1 [F_2,\ G_2] + G_1 [F_1,\ G_2] F_2$$
$$+ F_1 [F_2,\ G_1] G_2 + [F_1,\ G_1] G_2 F_2 \qquad (2.13^*)$$

Thus for $[F_1 F_2, G_1 G_2]$ we have two different expressions, which must be equal irrespective of the forms of F and G. Equating these two results, we have

$$(F_1 G_1 - G_1 F_1) [F_2,\ G_2] = [F_1,\ G_1] (F_2 G_2 - G_2 F_2) \qquad (2.14)$$

This condition holds only in one case: if for any two operators F and G

$$[F,\ G] = c\,(FG - GF) \qquad (2.15)$$

where c is an operator that commutes with any other operator. Only multiplication into a constant has such a property. Therefore c is a constant. We can easily show that c is pure imaginary. Indeed, we want the Poisson bracket of two real quantities to be real. Hence, if F and G are hermitian, $[F, G]$ must be hermitian too. But by Eq. (4.8), Chapter II,

$$[F,\ G]^+ = \bar{c}\,(G^+ F^+ - F^+ G^+) = -\,\bar{c}\,(FG - GF) \qquad (2.15^*)$$

If (2.15*) coincides with (2.15), then necessarily $\bar{c} = -c$, which yields

$$c = \frac{i}{h'} \qquad (2.16)$$

where h' is real. Thus

$$[F,\ G] = \frac{i}{h'}(FG - GF) \qquad (2.17)$$

Let us show that (2.17) satisfies all the properties of the Poisson bracket (2.7)-(2.11). We immediately see that (2.17) satisfies (2.7)-(2.9). Furthermore

$$\frac{i}{h'}(F_1 F_2 G - G F_1 F_2) = \frac{i}{h'} F_1 (F_2 G - G F_2) + \frac{i}{h'} (F_1 G - G F_1) F_2$$

which means that (2.10) is satisfied as well. Finally, to prove that (2.11) holds we must in the obvious identity

$$FGL + GLF + LFG + LGF + FLG + GFL$$
$$- FGL - GLF - LFG - LGF - FLG - GFL = 0$$

group the terms properly. We may thus consider it proved that the quantum Poisson bracket has the form (2.17). What remains to be found is the constant quantity h'. For the Poisson bracket to have the proper dimensions, h' must have the dimensions of action. The numerical value of h' can be found by first leaving it undefined, then constructing the appropriate operators, and finally comparing h' with the results of actual experiments, for example, comparing the eigenvalues of the energy operator with observed energy levels. We would then find that

$$h' = \hbar = \frac{h}{2\pi}, \qquad h = 6.624 \times 10^{-27} \text{ erg s} \qquad (2.18)$$

where h is Planck's constant. We will from the start think of h' as equal to $\hbar$.

Knowing the quantum Poisson bracket enables us to use the method of classical analogy to find the form of quantum operators. But, of course, the extent to which this analogy is justified will be determined by comparing theory with experiment.

3. Position and momentum operators

The form of the operator for a given physical quantity depends on the choice of variables for the functions on which the operator acts. The operator for the independent variable is always the multiplication into this variable. This follows from the requirement that the values of any physical quantity coincide with the eigenvalues of the corresponding operator (see Section 6 of Chapter II).

Let us take, for example, the coordinate x as the independent variable for a system with one degree of freedom. If operators act on the functions of x, the *position operator* will reduce to the operator corresponding to multiplication into x. If we take another quantity for the independent variable, for instance, energy, the energy operator will have the "multiplication" form, whereas the position operator will have a new, more complicated form. [2]

How does one choose the independent variables for a system with several degrees of freedom? Can any combination of variables be used (for instance, energy, one coordinate, and one component of the momentum vector for a system with three degrees of freedom)? We can answer these questions by reasoning as follows. The *independent-variable operators* are multiplication operators and thus commute with each other. But this means that

The independent variables are those whose operators commute.

---

[2] See Section 6, Chapter I, Part II.

To decide which quantities (operators) commute and which do not, we use the analogy between the classical and quantum Poisson brackets.

The electron of classical physics is a mass point with three degrees of freedom. Let us denote the position of an electron by the (Cartesian) coordinates

$$x = x_1, \qquad y = x_2, \qquad z = x_3 \tag{3.1}$$

and the momentum components by

$$p_x = p_1, \qquad p_y = p_2, \qquad p_z = p_3 \tag{3.2}$$

The classical Poisson brackets of these quantities are

$$[x_k, x_l] = 0, \qquad [p_k, p_l] = 0, \qquad [p_k, x_l] = \delta_{kl} \tag{3.3}$$

Let us try to formulate this in quantum terms. We will assume that the quantum Poisson brackets have the same form as (3.3) and that the 0 and 1 in the right-hand sides of (3.3) are the operators of multiplication into zero and unity respectively. If, in addition, we assume that the coordinates of the electron can take on any real values from $-\infty$ to $+\infty$, Eqs. (3.3) enable us to find the form of the position and momentum operators.

First of all, Eqs. (3.3) show that the operators for x_1, x_2, x_3 commute. Hence we must take x_1, x_2, x_3 for the independent variables. This means that the operators will act on functions of the type

$$\psi(x, y, z) = \psi(x_1, x_2, x_3) \tag{3.4}$$

Second, Eqs. (3.3) yield

$$\frac{i}{\hbar}(p_x x - x p_x)\psi = \psi$$

$$\frac{i}{\hbar}(p_y y - y p_y)\psi = \psi$$

$$\frac{i}{\hbar}(p_z z - z p_z)\psi = \psi \tag{3.5}$$

The *momentum operators*

$$p_x = -i\hbar\frac{\partial}{\partial x}, \qquad p_y = -i\hbar\frac{\partial}{\partial y}, \qquad p_z = -i\hbar\frac{\partial}{\partial z} \tag{3.6}$$

are, as we already know, hermitian. They also satisfy Eq. (3.5), since after cancelling out $\hbar$ we have

$$[p_x, x]\psi = \frac{\partial}{\partial x}(x\psi) - x\frac{\partial\psi}{\partial x} = \psi \tag{3.7}$$

We can obtain equations for the y and z components in the same way.

To find the general form of the momentum operators p_x, p_y, p_z we put

$$p'_x = -i\hbar \frac{\partial}{\partial x} + q_x, \quad p'_y = -i\hbar \frac{\partial}{\partial y} + q_y, \quad p'_z = -i\hbar \frac{\partial}{\partial z} + q_z \quad (3.8)$$

Equations (3.3) yield

$$q_k x_l - x_l q_k = 0, \qquad k, \, l = 1, 2, 3 \qquad (3.9)$$

which means that $q_x, \, q_y, q_z$ and x, y, z commute. Furthermore, the q's must be hermitian. Hence they correspond to multiplication into real functions of x, y, z. For example,

$$p'_x \psi' = -i\hbar \frac{\partial \psi'}{\partial x} + q_x (x, \, y, \, z) \, \psi' \qquad (3.10)$$

But $p'_x, \, p'_y, \, p'_z$ must also satisfy the conditions

$$[p'_k, \, p'_l] = \frac{i}{\hbar} (p'_k p'_l - p'_l p'_k) = 0$$

or

$$\frac{\partial}{\partial x_k} (q_l \psi') + q_k \frac{\partial \psi'}{\partial x_l} - \frac{\partial}{\partial x_l} (q_k \psi') - q_l \frac{\partial \psi'}{\partial x_k} = 0$$

that is

$$\left(\frac{\partial q_l}{\partial x_k} - \frac{\partial q_k}{\partial x_l} \right) \psi' = 0$$

which yields

$$\frac{\partial q_l}{\partial x_k} - \frac{\partial q_k}{\partial x_l} = 0 \qquad (3.11)$$

Hence q_x, q_y, q_z are partial derivatives of one real function of coordinates, with the result that

$$p'_x \psi' = -i\hbar \frac{\partial \psi'}{\partial x} + \frac{\partial f}{\partial x} \, \psi'$$

$$p'_y \psi' = -i\hbar \frac{\partial \psi'}{\partial y} + \frac{\partial f}{\partial y} \, \psi'$$

$$p'_z \psi' = -i\hbar \frac{\partial \psi'}{\partial z} + \frac{\partial f}{\partial z} \, \psi' \qquad (3.12)$$

We now show that by transforming the function ψ we can reduce the operators in (3.12) to the more simple form (3.6). Let us assume that

$$\psi' = e^{-if(x, \, y, \, z)/\hbar} \psi \qquad (3.13)$$

$$p^*_x \psi = e^{if/\hbar} p'_x \psi' = e^{if/\hbar} p'_x e^{-if/\hbar} \psi \qquad (3.14)$$

We seek the form of p_x^*. We have

$$- i\hbar \, \frac{\partial \psi}{\partial x} = e^{if/\hbar} \left(- i\hbar \frac{\partial \psi'}{\partial x} + \frac{\partial f}{\partial x} \, \psi' \right) = e^{if/\hbar} p_x' \psi' \qquad (3.15)$$

If we compare (3.14) and (3.15), we obtain

$$p_x^* \psi = - i\hbar \, \frac{\partial \psi}{\partial x}$$

and similar equations for y and z. We see that

$$p_x^* = p_x, \qquad p_y^* = p_y, \qquad p_z^* = p_z \qquad (3.16)$$

Thus the new operators are of form (3.6), and the relationship between p_x, p_y, p_z and p_x', p_y', p_z' is as follows:

$$p_x = e^{if/\hbar} p_x' e^{-if/\hbar}$$
$$p_y = e^{if/\hbar} p_y' e^{-if/\hbar}$$
$$p_z = e^{if/\hbar} p_z' e^{-if/\hbar} \qquad (3.17)$$

We will assume that this transformation was done at the very beginning. So we can view momentum operators p_x, p_y, p_z acting on functions of coordinates as the operators (3.6).

It is a distinctive feature of quantum mechanics that an operator for a physical quantity can have different forms [Eqs. (3.6) and (3.12), for instance] and that in passing from one form to another [Eqs. (3.17)] the wave function ψ undergoes a transformation [see Eq. (3.13)]. One might think that such arbitrariness would lead to an ambiguity in the laws of quantum mechanics. This is a false assumption. All quantities that can be compared with experimental results (eigenvalues of operators, for instance) are unambiguous. In Section 12 we will return to this question and find the transformations under which physical quantities are invariant.

4. Eigenfunctions and eigenvalues of the momentum operator

We see that using the method of classical analogy we can find the operators of the Cartesian components of momentum, which are

$$p_x = - i\hbar \frac{\partial}{\partial x}, \qquad p_y = - i\hbar \frac{\partial}{\partial y}, \qquad p_z = - i\hbar \frac{\partial}{\partial z} \qquad (4.1)$$

What are the eigenvalues and eigenfunctions of these operators? To solve this problem, we first denote the eigenvalues of

the corresponding operators as p'_x, p'_y, p'_z. The eigenvalue equations are

$$-i\hbar \frac{\partial \psi^{(1)}}{\partial x} = p'_x \psi^{(1)}$$

$$-i\hbar \frac{\partial \psi^{(2)}}{\partial y} = p'_y \psi^{(2)}$$

$$-i\hbar \frac{\partial \psi^{(3)}}{\partial z} = p'_z \psi^{(3)} \qquad (4.2)$$

The solution for the first of these equations is

$$\psi^{(1)} = f^{(1)}(y, z) e^{ixp'_x/\hbar} \qquad (4.3)$$

where $f^{(1)}$ does not depend on x. The solution is finite for all values of x if and only if p'_x is real. Thus the eigenvalues of p_x form a continuous spectrum ranging from $-\infty$ to $+\infty$. In a similar manner we can write the other two solutions:

$$\psi^{(2)} = f^{(2)}(z, x) e^{iyp'_y/\hbar}, \qquad \psi^{(3)} = f^{(3)}(x, y) e^{izp'_z/\hbar} \qquad (4.3^*)$$

It is easy to see that Eqs. (4.2) have the general solution

$$\psi^{(1)} = \psi^{(2)} = \psi^{(3)} = \psi \qquad (4.4)$$

where $\psi = c \exp[(i/\hbar)(xp'_x + yp'_y + zp'_z)]$ with c a constant that may, however, depend on p'_x, p'_y, p'_z. The normalization condition will determine this constant. We will first consider the one-dimensional case. We put

$$\psi = c e^{ixp/\hbar} \qquad (4.5)$$

The normalization condition [see Eq. (7.20), Chapter II] for this case is

$$\lim_{\Delta p \to 0} \frac{1}{\Delta p} \int_{-\infty}^{+\infty} |\Delta \Psi|^2 \, dx = 1 \qquad (4.6)$$

where

$$\Delta \Psi = \int_{p}^{p+\Delta p} \psi(x, p) \, dp \qquad (4.7)$$

Equation (4.6) is satisfied if we consider c not depending on p. We then obtain

$$\Delta \Psi = \frac{c\hbar}{ix} e^{ixp/\hbar} (e^{ix \, \Delta p/\hbar} - 1)$$

$$= c e^{ix \, (p+\Delta p/2)/\hbar} \frac{2\hbar}{x} \sin\left(\frac{x \, \Delta p}{2\hbar}\right) \qquad (4.8)$$

Next

$$\frac{1}{\Delta p} \int\limits_{-\infty}^{+\infty} |\Delta\Psi|^2 \, dx = \frac{4\,|c\,|^2\,\hbar^2}{\Delta p} \int\limits_{-\infty}^{+\infty} \sin^2\left(\frac{x\,\Delta p}{2\hbar}\right)\frac{dx}{x^2}$$

$$= 2\hbar\,|\,c\,|^2 \int\limits_{-\infty}^{+\infty} \frac{\sin^2 \xi}{\xi^2}\,d\xi = 2\pi\hbar\,|\,c\,|^2 \qquad (4.9)$$

since

$$\int\limits_{-\infty}^{+\infty} \frac{\sin^2 \xi}{\xi^2}\,d\xi = \pi \qquad (4.10)$$

Returning to (4.6), we find that

$$c = \frac{1}{(2\pi\hbar)^{1/2}}\,e^{i\alpha} \qquad (4.11)$$

where α is a real quantity, which we may consider equal to zero without loss of generality.

Hence the normalized function $\psi(x, p)$ is

$$\psi(x,\,p) = \frac{1}{(2\pi\hbar)^{1/2}}\,e^{ixp/\hbar} \qquad (4.12)$$

The expansion of an arbitrary function in momentum eigenfunctions will then be [according to Eqs. (8.21), Chapter II]

$$f(x) = \frac{1}{(2\pi\hbar)^{1/2}} \int\limits_{-\infty}^{+\infty} e^{ixp/\hbar}\varphi(p)\,dp \qquad (4.13)$$

where

$$\varphi(p) = \frac{1}{(2\pi\hbar)^{1/2}} \int\limits_{-\infty}^{+\infty} e^{-ixp/\hbar}f(x)\,dx \qquad (4.13^*)$$

Equations (4.13) and (4.13*) define the *Fourier integral transformations*, and $\varphi(p)$ is called the *Fourier transform of* $f(x)$ [quite naturally, $f(x)$ is called the *inverse Fourier transform* of $\varphi(p)$].

If we now consider the three-dimensional case, we must write the normalization condition as

$$\lim \frac{1}{\Delta p'_x\,\Delta p'_y\,\Delta p'_z} \iiint |\Delta\Psi|^2\,dx\,dy\,dz = 1 \qquad (4.14)$$

where

$$\Delta\Psi = \int\limits_{p'_x}^{p'_x + \Delta p'_x} dp'_x \int\limits_{p'_y}^{p'_y + \Delta p'_y} dp'_y \int\limits_{p'_z}^{p'_z + \Delta p'_z} dp'_z\,\psi\,(x,\,y,\,z;\,p'_x,\,p'_y,\,p'_z) \qquad (4.15)$$

The normalization condition (4.14) will obviously be satisfied if

$$\psi\left(x,\,y,\,z;\,p'_x,\,p'_y,\,p'_z\right) = \psi\left(x;\,p'_x\right)\psi\left(y;\,p'_y\right)\psi\left(z;\,p'_z\right)$$

$$= \frac{1}{(2\pi\hbar)^{3/2}} \exp\left[\frac{i}{\hbar}\left(xp'_x + yp'_y + zp'_z\right)\right] \quad (4.16)$$

The expansion of an arbitrary function of the three coordinates in the momentum eigenfunctions will then be written in the form of a triple integral.

5. Quantum description of systems

In Section 5, Chapter II, we found that an eigenvalue of an operator is related to the corresponding eigenfunction by the formula

$$\lambda = \frac{\int \bar{\psi} L \psi \, d\tau}{\int \bar{\psi} \psi \, d\tau} \quad (5.1)$$

We note that when ψ belongs to a continuous spectrum, we must take the proper differential instead of the eigenfunction. Thus by defining the eigenfunction we can find the corresponding eigenvalue of the operator, which in turn is related to the physical quantity. In this sense we can speak of a function as defining the state of the system.

The right-hand side of (5.1) retains its meaning when ψ is not an eigenfunction of L. We will clarify the physical meaning of this fact in the next chapter.

We will call a *wave function* a function ψ that defines the state of the system. Let us illustrate the important concept of a wave function as a means of describing a system's state. As we have shown in the previous section the function

$$\psi\left(x,\,y,\,z\right) = \frac{1}{(2\pi\hbar)^{3/2}} \exp\left[\frac{i}{\hbar}\left(xp'_x + yp'_y + zp'_z\right)\right] \quad (5.2)$$

is the simultaneous eigenfunction for all three momentum components p_x, p_y, p_z. It therefore defines the state of the electron with

$$p_x = p'_x, \qquad p_y = p'_y, \qquad p_z = p'_z \quad (5.3)$$

The other quantities (for instance, the coordinates of the electron) do not have definite values in the state of the electron with a wave function (5.2) since ψ is not an eigenfunction of the position operator.

Hence the quantum description of the state of the electron has the feature that only one group of quantities (for instance,

p_x, p_y, p_z) can have definite values; the other group (in our example x, y, z) remains undetermined. This result agrees with the fact mentioned in Chapter I, the impossibility of precise and simultaneous measurement of all the quantities that in classical physics characterize the state of the electron.

But how does one decide which quantities can be measured simultaneously and which cannot? Let us reason in the following manner. From the experiment we gain a knowledge of the state of the electron, that is, of a certain wave function ψ. If the measurement of two quantities L and M gives definite values λ and μ, then, according to what has just been said, ψ will be an eigenfunction of both L (with eigenvalue λ) and M (with eigenvalue μ). But for L and M to have simultaneous eigenfunctions both must satisfy certain conditions, which we will now examine.

6. Commutativity of operators

Let $\psi = \psi(x; \lambda, \mu)$ be a simultaneous eigenfunction [3] of operators L and M:

$$L\psi = \lambda\psi$$
$$M\psi = \mu\psi \qquad (6.1)$$

We operate on the first equation with M and on the second with L:

$$ML\psi = \lambda M\psi = \lambda\mu\psi$$
$$LM\psi = \mu L\psi = \mu\lambda\psi$$

Hence

$$ML\psi(x; \lambda, \mu) = LM\psi(x; \lambda, \mu) \qquad (6.2)$$

We assume now that all the simultaneous eigenfunctions constitute a complete set. Then any other function $\psi(x)$ can be expressed as a series (or an integral) of the type

$$\psi(x) = \sum_{\lambda, \mu} c(\lambda, \mu)\,\psi(x; \lambda, \mu) \qquad (6.3)$$

Since (6.2) holds for each term in expansion (6.3), it holds for the sum as a whole provided the series for $ML\psi$ converges. We see that for an arbitrary $\psi(x)$

$$ML\psi = LM\psi \qquad (6.4)$$

or

$$ML - LM = 0 \qquad (6.4^*)$$

---

[3] Letter x denotes either the independent variable or the set of all the independent variables of a function.

that is, L and M commute. We have thus proved the following theorem.

If the simultaneous eigenfunctions of two operators L and M form a complete set, the operators commute.

We will now prove the converse theorem:

If two operators L and M commute, they have simultaneous eigenfunctions.

Let eigenvalue λ of operator L have corresponding to it one or several eigenfunctions $\psi(x; \lambda, k)$, where k denotes the different functions for one specific value of λ. The most general solution of the eigenvalue problem

$$L\psi = \lambda\psi \tag{6.5}$$

is

$$\psi = \sum_k c(k)\, \psi(x; \lambda, k) \tag{6.6}$$

Operating with M on the equation

$$L\psi(x; \lambda, k) = \lambda\psi(x; \lambda, k) \tag{6.5*}$$

and using the fact that L and M commute, we obtain

$$ML\psi = L(M\psi) = \lambda M\psi \tag{6.7}$$

The function $\psi' = M\psi$ is thus the eigenfunction of L corresponding to the eigenvalue λ. Hence it can be expressed as a linear superposition of eigenfunctions $\psi(x; \lambda, k')$:

$$M\psi(x; \lambda, k) = \sum_{k'} M(k', k)\, \psi(x; \lambda, k') \tag{6.8}$$

where obviously the expansion coefficients $M(k', k)$ depend on k' as well as on k. With this in mind we can construct the linear superposition of type (6.6) that will also satisfy the equation

$$M\psi = \mu\psi \tag{6.9}$$

We substitute (6.6) into (6.9) and use Eq. (6.8). Equating the coefficients of $\psi(x; \lambda, k)$ yields a set of s equations

$$\sum_{k'} M(k, k')\, c(k') = \mu c(k) \tag{6.10}$$

where s is the number of possible values of k for a given λ. In other words, s is equal to the degree of degeneracy of λ. If we denote the solutions of Eqs. (6.10) as

$$c^{(1)}(k), \quad c^{(2)}(k), \ldots, c^{(s)}(k) \tag{6.11}$$

and the corresponding values of μ as

$$\mu_1, \quad \mu_2, \ldots, \mu_s \tag{6.12}$$

then the functions

$$\psi^* (x; \lambda, \mu_s) = \sum_k c^{(s)} (k) \psi (x; \lambda, k) \qquad (6.13)$$

will be the solutions to both (6.5) and (6.9), which means that they are simultaneous eigenfunctions of operators L and M. The proof is complete.

We assumed that λ was s-fold degenerate with s finite. This, in turn, leads to a finite number of eigenfunctions corresponding to each value of λ. The theorem also holds for the case when s is infinite.

What is the physical meaning of these theorems? It is the following:

The commutativity of operators expresses the possibility of measuring the corresponding physical quantities simultaneously. Or, in other words, the noncommutativity of operators makes a simultaneous measurement of the corresponding quantities impossible.

An example of two commutative operators with simultaneous eigenfunctions was given in Section 4 of this chapter.

7. Angular momentum

As an example of noncommutative operators let us examine three operators

$$m_x = yp_z - zp_y$$
$$m_y = zp_x - xp_z$$
$$m_z = xp_y - yp_x \qquad (7.1)$$

These are composed of position and momentum operators in the same way as angular momentum is in classical mechanics. Later in examining the quantum equations of the motion of the electron we will see that (7.1) can indeed be interpreted as *angular momentum operators*.

We set up the quantum Poisson bracket involving the operators (7.1) and the position and momentum operators. We find that

$$[m_x, x] = \frac{i}{\hbar} (m_x x - x m_x) = 0 \qquad (7.2)$$

This is obvious since m_x does not contain differentiation with respect to x and thus commutes with multiplication into x. We also find that

$$[m_x, y] = - z [p_y, y] = - z$$
$$[m_x, z] = y [p_z, z] \quad = y \qquad (7.2^*)$$

where we have used the properties of the Poisson bracket known from Section 2. In a similar manner

$$[m_x, p_x] = 0$$
$$[m_x, p_y] = -p_z$$
$$[m_x, p_z] = p_y \tag{7.3}$$

If we use (7.2) and (7.3) to find the Poisson bracket for two different components of angular momentum, we obtain

$$[m_x, m_y] = [m_x, zp_x - xp_z]$$
$$= [m_x, z] p_x - x [m_x, p_z]$$
$$= yp_x - xp_y$$
$$= -m_z$$

Hence

$$[m_y, m_z] = -m_x$$
$$[m_z, m_x] = -m_y$$
$$[m_x, m_y] = -m_z \tag{7.4}$$

We see that all these relationships correspond exactly to the classical ones.

Our task now is to find the eigenvalues and eigenfunctions of the operators m_x, m_y, m_z. The eigenvalue equation for m_z iis

$$\frac{\hbar}{i} \left(x \frac{\partial \psi}{\partial y} - y \frac{\partial \psi}{\partial x} \right) = m_z' \psi \tag{7.5}$$

with m_z' being an eigenvalue of m_z. If we use cylindrical coordinates ρ, φ, z with

$$x = \rho \cos \varphi, \qquad y = \rho \sin \varphi$$

then Eq. (7.5) reads

$$\frac{\hbar}{i} \frac{\partial \psi}{\partial \varphi} = m_z' \psi \tag{7.5*}$$

The solution is of the form

$$\psi = \psi^0 (z, \rho) e^{im_z' \varphi / \hbar} \tag{7.6}$$

This is a single-valued function of a point in space only if it is a periodic function of φ with period 2π, which yields

$$m_z' = m_3 \hbar, \qquad m_3 = 0, \pm 1, \pm 2, \ldots \tag{7.7}$$

We have thus found the eigenvalues and eigenfunctions of m_z. In the same way one can find them for the other two operators. To compare them we will return to the rectangular Cartesian coordi-

nates. Function (7.6), which we denote ψ_3, and the eigenfunctions ψ_1 and ψ_2 of m_x and m_y will then be of the form

$$\psi_1 = f_1\left(x,\ \sqrt{y^2 + z^2}\right)(y + iz)^{m_1}$$

$$\psi_2 = f_2\left(y,\ \sqrt{z^2 + x^2}\right)(z + ix)^{m_2}$$

$$\psi_3 = f_3\left(z,\ \sqrt{x^2 + y^2}\right)(x + iy)^{m_3} \qquad (7.8)$$

and the eigenvalues will be

$$m'_x = m_1\hbar, \qquad m'_y = m_2\hbar, \qquad m'_z = m_3\hbar \qquad (7.9)$$

where m_1, m_2, m_3 are integers.

We have come to a conclusion that at first glance seems paradoxical: the angular momentum components in any direction can assume, when measured, only integral values that are multiples of a definite quantity $\hbar$. This seems even stranger in view of the fact that the components of a vector in two infinitely close directions differ by an infinitely small quantity.

But this paradox is easily explained. First let us note that the only simultaneous eigenfunction of m_x, m_y, m_z corresponds to the simultaneous eigenvalues

$$m_1 = m_2 = m_3 = 0 \qquad (7.10)$$

and is

$$\psi = \psi_1 = \psi_2 = \psi_3 = f(r), \qquad r = (x^2 + y^2 + z^2)^{1/2} \qquad (7.11)$$

In this case the *angular momentum vector* (and hence its components in any direction) is zero, which implies that a paradox does not exist. But if one of the eigenvalues is nonzero, the operators m_x, m_y, m_z have not a single simultaneous eigenfunction. Hence a state of the electron in which two or three components of angular momentum have definite values simultaneously is impossible. This means that only one of the components can be integral. What is the physical significance of this fact? To measure a component of an electron's angular momentum in a definite direction one must influence the electron in some way, say, by switching on a magnetic field in this direction. This influence "tunes" the electron in such a way that its angular momentum component in the direction of the field assumes integral values. Other components remain undefinable because it is impossible to measure them without changing the direction of the field, that is, without getting the electron "out of tune". We conclude that the properties of angular momentum resulting from the theory under consideration express the inevitable influence of the measuring process on an object.

8. The energy operator

In the classical theory for a vast variety of systems the time dependence of the state of a system (equations of motion) is defined by introducing the Hamiltonian function, which is the total energy of the system. In quantum mechanics too we can introduce an energy operator, commonly called the *Hamiltonian*, which defines the time dependence of the state of a system (this will be proved in Section 13 of this chapter). For this reason the choice of a specific Hamiltonian is an important step in constructing the theory. When this is done, the choice of operators for the other physical quantities (angular momentum, for example) is more limited. In classical mechanics we take the simplest quantities (position and momentum) and construct different combinations possessing "advantageous" properties (for instance, remaining constant in the process of motion). In quantum mechanics too we take the simplest operators and construct combinations that have simple properties and allow an obvious interpretation. When speaking of the properties of operators, we have in mind for the most part their commutativity with other operators and their time dependence. Since the time dependence is associated with the form of the Hamiltonian (see Section 13 of this chapter), it is evident that the choice of "advantageous" combinations of the simplest operators depends on this basic operator.

The classical Hamiltonian function has different forms depending on whether special theory of relativity is taken into account. Only for the one-body problem was the relativistic Hamiltonian function found in explicit form. For the many-body problem this proved to be impossible. The situation is similar in quantum mechanics. Here too the relativistic Hamiltonian was found only for the one-body problem, and it differs drastically from the nonrelativistic Hamiltonian. We will study the relativistic case in Part V, which is devoted to Dirac's theory of the electron. For the present we will deal with the Hamiltonian ignoring the relativistic effects.

In the classical theory the kinetic energy expressed in terms of the components of momentum in a rectangular Cartesian coordinate system is

$$T = \frac{1}{2m} (p_x^2 + p_y^2 + p_z^2) \tag{8.1}$$

If we consider p_x, p_y, p_z as operators (3.6), formula (8.1) will become an operator, which we can call the *kinetic-energy operator*. We note that if in (8.1) we had used, say, a spherical coordinate system instead of a Cartesian one, then

$$T^* = \frac{1}{2m} \left(p_r^2 + \frac{1}{r^2} p_\theta^2 + \frac{1}{r^2 \sin^2 \theta} p_\varphi^2 \right) \tag{8.2}$$

If we had then interpreted p_r, p_θ, p_φ as $-i\hbar(\partial/\partial r)$, $-i\hbar(\partial/\partial\theta)$, $-i\hbar(\partial/\partial\varphi)$, we would have obtained an operator T^* that would not coincide with T. For this reason let us assume that when passing from a classical function to a quantum mechanical operator we must use only rectangular coordinates. If a classical formula (involving components in a rectangular coordinate system) does not contain factors that become noncommutative when transformed into operators, the transformation is unique. But, of course, it remains to be proved by comparing theory with experimental practice whether such classical analogy is legitimate.

The kinetic-energy operator can be expressed in terms of the Laplacian operator. If we substitute (3.6) for p_x, p_y, p_z, we obtain

$$T\psi = -\frac{\hbar^2}{2m}\nabla^2\psi \tag{8.3}$$

Naturally, after the form of the operator is established, we can transform to any coordinates. For instance, in spherical coordinates

$$T\psi = -\frac{\hbar^2}{2m}\left\{\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\psi}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right) \right.$$
$$\left. + \frac{1}{r^2\sin^2\theta}\frac{\partial^2\psi}{\partial\varphi^2}\right\} \tag{8.4}$$

If we introduce the operators

$$p_r = -i\hbar\frac{\partial}{\partial r}, \qquad p_\theta = -i\hbar\frac{\partial}{\partial\theta}, \qquad p_\varphi = -i\hbar\frac{\partial}{\partial\varphi} \tag{8.5}$$

we can write the kinetic-energy operator as

$$T = \frac{1}{2m}\left\{\frac{1}{r^2}p_r r^2 p_r + \frac{1}{r^2\sin\theta}p_\theta\sin\theta\, p_\theta + \frac{1}{r^2\sin^2\theta}p_\varphi^2\right\} \tag{8.6}$$

This expression differs from (8.2) only in the order of the noncommutative multipliers; if they were commutative, the two would have coincided.

We know that the eigenvalues of the Laplacian operator are negative. Hence the eigenvalues of the kinetic-energy operator are positive, as they should be.

For the eigenfunctions of the kinetic-energy operator we can take the simultaneous eigenfunctions of p_x, p_y, p_z, which, as we know, are of the form

$$\psi = \frac{1}{(2\pi\hbar)^{3/2}}\exp\left[\frac{i}{\hbar}\left(xp_x' + yp_y' + zp_z'\right)\right] \tag{8.7}$$

Any function of type (8.7) in which the sum of the squares of the parameters p_x', p_y', p_z' has a definite value $2mT'$:

$$p_x'^2 + p_y'^2 + p_z'^2 = 2mT' \tag{8.8}$$

and also any superposition of such functions (a sum or an integral) is an eigenfunction of the kinetic-energy operator corresponding to the eigenvalue $T = T'$. Hence we have an eigenvalue of infinite-fold degeneracy. Functions of type (8.7) can be used to compose superpositions that would at the same time be eigenfunctions of other operators. This implies that these operators commute with each other and with the kinetic-energy operator. In the physical sense this means that the state of the electron is not fully defined when we specify the kinetic energy alone. Hence we must indicate values of other quantities, say, momentum.

The kinetic-energy operator for a free electron is at the same time its Hamiltonian. For an electron in an external field with a potential energy $U(x, y, z)$ we can by analogy with the classical theory write the Hamiltonian in the form of a sum of the kinetic- and *potential-energy* operators:

$$H = \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2) + U(x, y, z) \qquad (8.9)$$

Here the operator $U(x, y, z)$ acts on functions of coordinates and thus yields multiplication into $U(x, y, z)$. The eigenvalue equation for H is

$$-\frac{\hbar^2}{2m}\nabla^2\psi + U(x, y, z)\psi = E\psi \qquad (8.10)$$

where E is the total energy of the electron in the field. This equation was proposed in 1926 by Schrödinger and is called the *Schrödinger equation*. We will examine the equation and its solutions more closely in Part II. For the present we will say that, apart from certain details, its corollaries are borne out by experiments, which proves the validity of the original hypotheses.

The Schrödinger equation can serve to describe the behaviour of an electron in an electrostatic field. It is only natural then to try to generalize it for the case of a magnetic field. But it appears that the classical model of the electron as an electrically charged mass point is not sufficient to explain the electron's behaviour in a magnetic field and it is necessary to assign a definite magnetic moment to the electron. In Part V we will perform this generalization on the basis of Dirac's theory of the electron.

9. Canonical transformation

We have seen that the state of the electron can be described by a function of coordinates or other independent variables, for instance, the components of momentum. The transition from one set of independent variables to another is done by *canonical transformation*.

Let a wave function expressed in terms of coordinates be $\psi(x, y, z)$ or simply $\psi(x)$ if by x we denote the set of all three coordinates. We assume that we are dealing with an operator L that has eigenvalues λ and a complete set of eigenfunctions $\varphi(x, \lambda)$. Then $\psi(x)$ can be expanded in a series in the eigenfunctions $\varphi(x, \lambda)$:

$$\psi(x) = \sum_k c(\lambda_k) \varphi(x, \lambda_k) + \int c(\lambda) \varphi(x, \lambda) \, d\lambda \qquad (9.1)$$

where the expansion coefficients [both $c(\lambda_k)$ and $c(\lambda)$] are defined in terms of $\psi(x)$ thus:

$$c(\lambda) = \int \overline{\varphi(x, \lambda)} \, \psi(x) \, dx \qquad (9.2)$$

Function $\psi(x)$ is specified by the set of coefficients $c(\lambda_k)$ and $c(\lambda)$. For this reason, if $\psi(x)$ describes the state in terms of x, then $c(\lambda)$ describes it in terms of λ. Also, if $\psi(x)$ is normalized, so is $c(\lambda)$ because according to the completeness condition (see Section 8 of the previous chapter)

$$\int |\psi(x)|^2 \, dx = \sum_k |c(\lambda_k)|^2 + \int |c(\lambda)|^2 \, d\lambda \qquad (9.3)$$

A state in which $\lambda = \lambda_n$ is described in terms of λ by the function

$$c(\lambda_n) = 1$$
$$c(\lambda) = 0, \qquad \lambda \neq \lambda_n \qquad (9.4)$$

If in a given state we find that $\lambda = \lambda'$ where λ' belongs to a continuous spectrum, the expansion coefficients $c(\lambda_k)$ in (9.1) must be set equal to zero and the integral must be the Stieltjes integral. We write it in the form

$$\psi(x) = \varphi(x, \lambda') = \int \varphi(x, \lambda) \, d_\lambda c(\lambda, \lambda') \qquad (9.5)$$

where

$$c(\lambda, \lambda') = 1, \qquad \lambda > \lambda'$$
$$c(\lambda, \lambda') = 0, \qquad \lambda \leqslant \lambda' \qquad (9.6)$$

Formula (9.2) can be considered to be the expansion of $c(\lambda)$ in the functions

$$\varphi^+(\lambda, x) = \overline{\varphi(x, \lambda)} \qquad (9.7)$$

Here the expansion coefficients are the functions $\psi(x)$ defined by (9.1). Later we will see that the φ^+ are the eigenfunctions of the operator x in terms of λ. Thus a description in terms of x is equivalent to a description in terms of λ.

Now let us see how operators change when we transfer from one set of variables to another. We first take the operator L whose eigenfunctions we have used in expansions. We apply it to $\psi(x)$. Since $\varphi(x, \lambda)$ is an eigenfunction of L,

$$L\varphi(x, \lambda) = \lambda\varphi(x, \lambda) \tag{9.8}$$

and we obtain

$$L\psi(x) = \sum_k \lambda_k c(\lambda_k) \varphi(x, \lambda_k) + \int \lambda c(\lambda) \varphi(x, \lambda)\, d\lambda \tag{9.9}$$

Hence the transition from $\psi(x)$ to $L\psi(x)$, that is the application of L, has corresponding to it the transition from $c(\lambda)$ to $\lambda c(\lambda)$, or the multiplication into λ. Consequently, the operator L expressed in terms of the independent variables λ, its eigenvalues, is simply the multiplication into λ, which is as it should be. Indeed, we have seen in Section 3 that the operator corresponding to the independent variable is the multiplication into this variable.

Now instead of L let us take another operator M and apply it to $\psi(x)$. For simplicity let us assume that L has only a discrete spectrum. So the series expansion of $\psi(x)$ in the eigenfunctions of L will be

$$\psi(x) = \sum_k c(\lambda_k) \varphi(x, \lambda_k) \tag{9.10}$$

Applying M to $\psi(x)$, we find that

$$M\psi(x) = \sum_k c(\lambda_k) M\varphi(x, \lambda_k) \tag{9.11}$$

In turn we expand $M\varphi(x, \lambda_k)$ in the $\varphi(x, \lambda_n)$:

$$M\varphi(x, \lambda_k) = \sum_n (\lambda_n | M | \lambda_k) \varphi(x, \lambda_n) \tag{9.12}$$

where $(\lambda_n | M | \lambda_k)$ are the expansion coefficients:

$$(\lambda_n | M | \lambda_k) = \int \overline{\varphi(x, \lambda_n)} M\varphi(x, \lambda_k)\, dx \tag{9.13}$$

Substituting (9.12) into (9.11), we obtain

$$M\psi(x) = \sum_n c'(\lambda_n) \varphi(x, \lambda_n) \tag{9.14}$$

with $c'(\lambda_n)$ defined as

$$c'(\lambda_n) = Mc(\lambda_n) = \sum_k (\lambda_n | M | \lambda_k) c(\lambda_k) \tag{9.15}$$

Hence the transition from $\psi(x)$ to $M\psi(x)$ has corresponding to it the transition from $c(\lambda_n)$ to $c'(\lambda_n) = Mc(\lambda_n)$. Therefore M expressed in terms of λ has the form (9.15).

If L has a continuous spectrum as well as a discrete spectrum, instead of (9.10), (9.12), (9.14), and (9.15) we have

$$\psi(x) = \sum_k c(\lambda_k) \varphi(x, \lambda_k) + \int c(\lambda) \varphi(x, \lambda) \, d\lambda \qquad (9.10^*)$$

$$M\varphi(x, \lambda) = \sum_n (\lambda_n | M | \lambda) \varphi(x, \lambda_n) + \int (\lambda' | M | \lambda) \varphi(x, \lambda') \, d\lambda' \quad (9.12^*)$$

$$M\psi(x) = \sum_k c'(\lambda_k) \varphi(x, \lambda_k) + \int c'(\lambda) \varphi(x, \lambda) \, d\lambda \qquad (9.14^*)$$

$$c'(\lambda) = Mc(\lambda) = \sum_k (\lambda | M | \lambda_k) c(\lambda_k) + \int (\lambda | M | \lambda') c(\lambda') \, d\lambda' \quad (9.15^*)$$

where in (9.12*) and (9.15*) λ is an eigenvalue belonging either to the continuous or to the discrete spectrum. We can define $(\lambda | M | \lambda')$ in the same way as in (9.13). It could happen, however, that the integral in the expression for $(\lambda | M | \lambda')$ does not have a definite value. This means that in terms of λ there is no kernel for M. In such a case we construe M in terms of λ as an operator that maps the expansion coefficients $c(\lambda)$ of $\psi(x)$ into the expansion coefficients $c'(\lambda)$ of $M\psi(x)$ even if $c'(\lambda)$ is not given by (9.15*).

Let us illustrate this. We assume that M is the position operator, which means that when M is applied to $\psi(x)$, the result is multiplication into x. What we want to find is the expression for x in terms of λ. The eigenvalue equation is

$$\sum_k (\lambda | x | \lambda_k) c(\lambda_k) + \int (\lambda | x | \lambda') c(\lambda') \, d\lambda' = xc(\lambda) \qquad (9.16)$$

It is easy to check that the solution is

$$c(\lambda) = \varphi^+(\lambda, x) = \overline{\varphi(x, \lambda)} \qquad (9.17)$$

with $\varphi(x, \lambda)$ being an eigenfunction of the operator L in terms of x. Indeed, if we recall the condition

$$(\lambda | x | \lambda') = \overline{(\lambda' | x | \lambda)} \qquad (9.18)$$

which expresses the hermiticity of the operator x, and if we replace $c(\lambda)$ by $\overline{\varphi(x, \lambda)}$, we see that the complex conjugate of Eq. (9.16) is

$$\sum_k (\lambda_k | x | \lambda) \varphi(x, \lambda_k) + \int (\lambda' | x | \lambda) \varphi(x, \lambda') \, d\lambda' = x\varphi(x, \lambda) \quad (9.19)$$

But this is precisely the expansion of the product $x\varphi(x, \lambda)$ in a complete set of eigenfunctions $\varphi(x, \lambda')$. This also holds when instead of x we take an arbitrary operator M. Thus

The eigenfunctions of an operator M in terms of L are complex conjugates to the eigenfunctions of L in terms of M.

The result remains valid when operators do not have kernels.

10. An example of canonical transformation

As an example of canonical transformations we will consider transformations of the momentum and position operators, p and x. We know that p in terms of x, that is,

$$p = - i\hbar \frac{\partial}{\partial x} \qquad (10.1)$$

has eigenfunctions of the form

$$\psi(x, p) = \frac{1}{(2\pi\hbar)^{1/2}} e^{ixp/\hbar} \qquad (10.2)$$

Let us examine the form of x in terms of p. The operator x, by definition, maps $f(p)$, which is the Fourier transform of the function $\psi(x)$ expressed as the Fourier integral

$$\psi(x) = \frac{1}{(2\pi\hbar)^{1/2}} \int_{-\infty}^{+\infty} f(p) e^{ixp/\hbar} \, dp \qquad (10.3)$$

into another function $f'(p)$ such that

$$x\psi(x) = \frac{1}{(2\pi\hbar)^{1/2}} \int_{-\infty}^{+\infty} f'(p) e^{ixp/\hbar} \, dp \qquad (10.4)$$

holds true. If we integrate by parts, we obtain

$$x\psi(x) = \frac{1}{(2\pi\hbar)^{1/2}} \frac{\hbar}{i} \int_{-\infty}^{+\infty} f(p) \, de^{ixp/\hbar} = \frac{1}{(2\pi\hbar)^{1/2}} \int_{-\infty}^{+\infty} i\hbar \frac{\partial f}{\partial p} e^{ixp/\hbar} \, dp$$

which yields

$$f'(p) = i\hbar \frac{\partial f}{\partial p} \qquad (10.5)$$

We see that x in terms of p is

$$x = i\hbar \frac{\partial}{\partial p} \qquad (10.6)$$

This agrees with the expression for the Poisson bracket

$$[p, x] = \frac{i}{\hbar} (px - xp) = 1 \qquad (10.7)$$

because

$$- p \frac{\partial f}{\partial p} + \frac{\partial\, pf}{\partial p} = f \tag{10.8}$$

The eigenfunctions of x in terms of p are

$$\psi^+ (p,\, x) = \frac{1}{(2\pi\hbar)^{1/2}} e^{-ixp/\hbar} \tag{10.9}$$

Here p stands for the independent variable and x for a parameter, whereas in (10.2) the roles of p and x are interchanged.

The function

$$\psi (x) = \frac{1}{(2\pi\hbar)^{1/2}} e^{ixp/\hbar} \tag{10.10}$$

of x described a state of the electron with a definite momentum, p. However, a state with a definite position, x', was described in terms of x by a proper differential

$$dF (x,\, x') \tag{10.11}$$

with

$$\begin{aligned} F (x,\, x') &= 1, \qquad x > x' \\ F (x,\, x') &= 0, \qquad x \leqslant x' \end{aligned} \tag{10.12}$$

On the other hand, in terms of p a state with a definite position, x, is described by

$$f (p) = \frac{1}{(2\pi\hbar)^{1/2}} e^{-ixp/\hbar} \tag{10.13}$$

and with a definite momentum, p', by a proper differential

$$dF (p,\, p') \tag{10.14}$$

We must note that $F(p, p')$ depends on p and p' in the same way as $F(x, x')$ depends on x and x'.

It can easily be concluded, therefore, that a transition from one *representation* to another is made, as in the general case, via the Fourier integral, because for a state of the electron with a definite position, $x = x'$,

$$f (p) = \psi^+ (p,\, x') = \int \psi^+ (p,\, x)\, d_x F (x,\, x') \tag{10.15}$$

and for a state with a definite momentum, $p = p'$,

$$\psi (x) = \psi (x,\, p') = \int \psi (x,\, p)\, d_p F (p,\, p') \tag{10.16}$$

11. Canonical transformation as an operator

A canonical transformation is most conveniently written in symbolic notation. We will denote by $S(x, \lambda)$ the operator that maps a function $c(\lambda)$, which describes a state in terms of λ, into

a function $\psi(x)$, which describes the same state in terms of x. We can then proceed to write the expansion (9.1) in symbolic form:

$$\psi(x) = S(x, \lambda) c(\lambda) \tag{11.1}$$

What distinguishes this new operator from all the operators that we have studied previously is that it transforms a function of a definite independent variable into a function of another independent variable, both functions describing the *same* state in terms of *different* variables.

We can write the dependence of $c(\lambda)$ on $\psi(x)$, Eq. (9.2), in the form

$$c(\lambda) = S^{-1}(\lambda, x) \psi(x) \tag{11.2}$$

Let us show that the inverse $S^{-1}(\lambda, x)$ coincides with the hermitian conjugate $S^+(\lambda, x)$. Together with the functions $\psi(x)$ and $c(\lambda)$ we will consider two other functions, $\psi'(x)$ and $c'(\lambda)$, also related by Eqs. (11.1) and (11.2). We generalize the previous definition of a hermitian conjugate for the case of two independent variables and denote $S^+(\lambda, x)$ as an operator that satisfies the condition

$$\int \overline{\psi'(x)} \left[S(x, \lambda) c(\lambda) \right] dx = \int \overline{\left[S^+(\lambda, x) \psi'(x) \right]} c(\lambda) d\lambda \tag{11.3}$$

(In the case of a discrete spectrum we substitute a sum for an integral.) By (11.1) and the completeness condition the left side of (11.3) is

$$\int \overline{\psi'(x)} \psi(x) dx = \int \overline{c'(\lambda)} c(\lambda) d\lambda \tag{11.4}$$

The right sides of (11.3) and (11.4) coincide for any $c(\lambda)$ only if

$$c'(\lambda) = S^+(\lambda, x) \psi'(x) \tag{11.5}$$

for any $\psi'(x)$. If we compare this with (11.2), we find that

$$S^{-1}(\lambda, x) = S^+(\lambda, x) \tag{11.6}$$

and consequently

$$S(x, \lambda) S^+(\lambda, x) = 1 \tag{11.7}$$

$$S^+(\lambda, x) S(x, \lambda) = 1 \tag{11.7*}$$

As we know, an operator that obeys these conditions is called unitary. Thus

A transformation from one set of variables to another is done by means of a unitary operator.

Let us see how we can use S to express a canonical transformation of an operator corresponding to a definite physical quan-

tity M. If M in terms of x maps $\psi(x)$ into $\psi'(x)$, that is, if

$$\psi'(x) = M(x)\psi(x) \tag{11.8}$$

then the same operator M in terms of λ will, as we already know, map $c(\lambda)$ into $c'(\lambda)$:

$$c'(\lambda) = M(\lambda)c(\lambda) \tag{11.9}$$

But

$$c'(\lambda) = S^+(\lambda, x)\psi'(x) = S^+(\lambda, x)M(x)\psi(x) \tag{11.10}$$

and

$$\psi(x) = S(x, \lambda)c(\lambda) \tag{11.11}$$

Substituting (11.11) into (11.10), we obtain

$$c'(\lambda) = S^+(\lambda, x)M(x)S(x, \lambda)c(\lambda) \tag{11.12}$$

which after comparison with (11.9) yields

$$M(\lambda) = S^+(\lambda, x)M(x)S(x, \lambda) \tag{11.13}$$

Hence the transformation (9.11) of $\psi(x)$ corresponds to the transformation (9.13) of $M(x)$.

Obviously, if two unitary transformations are applied in succession, the result is a third unitary transformation. Indeed, instead of first transforming from a λ-representation to an x-representation by means of a unitary transformation (operator) $S(x, \lambda)$ and then from the x-representation to a μ-representation by means of $T(\mu, x)$, we can transform directly from the λ- to the μ-representation by means of

$$U(\mu, \lambda) = T(\mu, x)S(x, \lambda) \tag{11.14}$$

which clearly is a unitary transformation (operator) too. Generally speaking, $S(x, \lambda)$ has a kernel. If we compare (11.1) with (9.1) and (11.2) or (11.5) with (9.2), we can easily see that

$$kernel\ \ S(x, \lambda) = \varphi(x, \lambda) \tag{11.15}$$

$$kernel\ \ S^+(\lambda, x) = \varphi^+(\lambda, x) = \overline{\varphi(x, \lambda)} \tag{11.15*}$$

The kernel of the operator of the unitary transformation from L-representation (that is, from λ-representation) to x-representation is an eigenfunction of the operator L in terms of x.

12. Unitary invariants

In the process of finding the momentum operator in Section 3 we encountered the following fact. Not only the operator

$$p_k\psi = -i\hbar\frac{\partial\psi}{\partial x_k}, \qquad k = 1, 2, 3 \tag{12.1}$$

but the operators

$$p'_k\psi' = -i\hbar\frac{\partial\psi'}{\partial x_k} + \frac{\partial f}{\partial x_k}\psi' \qquad (12.1^*)$$

as well obeyed all the conditions resulting from the form of the Poisson bracket. The functions ψ and ψ' were related thus:

$$\psi' = e^{-if/\hbar}\psi \qquad (12.2)$$

and the operators p_k and p'_k thus:

$$p'_k = e^{-if/\hbar}p_k e^{if/\hbar} \qquad (12.3)$$

The real function $f(x_1, x_2, x_3)$ remained undefined.

If we compare (12.2) with (11.11) and (12.3) with (11.13), we see that Eqs. (12.2) and (12.3) represent a unitary transformation of a special type. Namely, the transformation is not associated with a change in variables. It results in multiplication by a function of the independent variables, the function being of modulus unity. The operator of this transformation

$$S = e^{if/\hbar} \qquad (12.4)$$

is called a *phase factor*.

We know that an operator for a given physical quantity can be described in terms of different independent variables, or, so to say, be in different representations. Now even when the representation is chosen we are left with an arbitrary phase factor. Both the transformation from one representation to another and the introduction of a phase factor are determined by unitary transformation. This implies that any two representations of an operator are interrelated by means of a unitary transformation. We can say that

The form of an operator for a given physical quantity is determined by the properties of the quantity only up to a unitary transformation.

Since the properties of physical systems cannot have undefined elements, they must be expressed by mathematical relationships that remain invariant under unitary transformations. It is the invariants that play an important role in the theory.

What do we mean by *unitary invariants?* The spectrum of eigenvalues of an operator is one of such invariants. So is the hermiticity of an operator. Indeed, by Eq. (11.13),

$$M(\lambda) = S^+M(x)S \qquad (12.5)$$

Hence, according to the rule of finding the hermitian conjugate of a product of operators,

$$M^+ (\lambda) = S^+ M^+ (x) (S^+)^+ = S^+ M^+ (x) S \qquad (12.6)$$

We come to the conclusion that from

$$M^+ (x) = M (x)$$

follows

$$M^+ (\lambda) = M (\lambda)$$

Also, the equations

$$M (x) \psi (x) = \mu \psi (x) \qquad (12.7)$$

$$M (\lambda) c (\lambda) = \mu c (\lambda) \qquad (12.8)$$

are equivalent, since one can be obtained from the other by means of the transformation

$$\psi (x) = Sc (\lambda) \qquad (12.9)$$

For this reason the eigenvalues μ in (12.7) and (12.8) are the same. This fact is closely related to the completeness condition, according to which

$$\int \overline{\psi (x)} \, \psi' (x) \, dx = \int \overline{c (\lambda)} \, c' (\lambda) \, d\lambda \qquad (12.10)$$

for any two pairs of functions [$\psi(x)$ and $c(\lambda)$, and $\psi'(x)$ and $c'(\lambda)$] that satisfy Eq. (12.9). This leads us to believe that an integral of type (12.10) is a unitary invariant. So is the following integral:

$$\int \overline{\psi (x)} \, M (x) \, \psi (x) \, dx = \int \overline{c (\lambda)} \, M (\lambda) \, c (\lambda) \, d\lambda \qquad (12.11)$$

whose physical meaning will be discussed in the next chapter.

Finally, any algebraic equation between linear operators, for example,

$$N (x) = M (x) + L (x)$$

or

$$N (x) = M (x) L (x)$$

is left invariant by a unitary transformation, since if all three operators $L(x)$, $M(x)$, $N(x)$ are subjected to the same transformation, the new operators $L(\lambda)$, $M(\lambda)$, $N(\lambda)$ will be related by the same equations. For instance, in transforming the operators x and p_x to any new variables, the Poisson bracket

$$\frac{i}{\hbar} (p_x x - x p_x) = 1$$

remains equal to unity.

13. Time evolution of systems.
Time dependence of operators

When we considered the operators for different physical quantities we did not account for time development. Yet in classical mechanics all quantities depend on time. What is the equivalent of this concept in quantum mechanics?

We know that an operator for a given quantity allows for different mathematical representations, and the choice of a specific representation remains arbitrary. Let us consider a representation in which the mathematical form of the operators remains the same for all instances of time t. Such a representation is possible only if the eigenvalue spectrum of an operator does not change in time, as is usually the case. For instance, the momentum operator p_x can be represented as $-i\hbar(\partial/\partial x)$ for any t. If at time $t = 0$ the momentum had a definite eigenvalue, for instance, $p_x = p'_x$, so that

$$p_x \psi = p'_x \psi, \qquad t = 0 \tag{13.1}$$

then at time $t > 0$ it can, generally speaking, take on any other eigenvalue or become undefined. In our example

$$p_x \psi \neq p'_x \psi, \qquad t > 0 \tag{13.2}$$

Since the form of p_x is assumed to remain unaltered, it is the form of the wave function ψ that changes. Thus

If we choose a representation in which the mathematical form of the operators does not change with time, the state of the system must be described by a time-dependent wave function.

This *time evolution* can be symbolically written as

$$\psi(x, t) = S(t)\,\psi(x, 0) \tag{13.3}$$

with the *time evolution operator* $S(t)$ being a continuous function of the time variable t and turning into the identity operator at time zero:

$$S(0) = 1 \tag{13.4}$$

We will assume the time evolution operator to be unitary:

$$S^+(t)\,S(t) = 1, \qquad S(t)\,S^+(t) = 1 \tag{13.5}$$

so as to preserve the normalization properties of our wave function as time passes:

$$\int \overline{\psi(x, t)}\,\psi(x, t)\,d\tau = \int \overline{\psi(x, 0)}\,\psi(x, 0)\,d\tau \tag{13.6}$$

Let us now turn to another way of representing operators. We recall that if a function ψ is transformed by means of a unitary

transformation of type (13.3), all operators will undergo a unitary transformation

$$L'(t) = S^+(t) L S(t) \qquad (13.7)$$

and the equations

$$\psi'(x,\,t) = L\psi(x,\,t) \qquad (13.8)$$

$$\psi'(x) = L'(t)\,\psi(x) \qquad (13.8^*)$$

are equivalent if

$$\psi'(x) = S^+\psi'(x,\,t) \qquad (13.9)$$

$$\psi(x) = S^+\psi(x,\,t) \qquad (13.9^*)$$

This second representation of operators lies in the fact that the operator corresponding to the quantity L at time $t > 0$ will be a *time-varying operator* $L'(t)$ defined by (13.7).

The difference between the two representations corresponds to the difference in the ways of describing a system. In the first the state of the system is described by a wave function of coordinates and time, and in the second by a wave function of coordinates exclusively, with time entering, if at all, as a parameter. If the initial state (at $t = 0$) was

$$\psi(x) = \psi(x,\,0)$$

then in the first representation the state at time $t > 0$ will be

$$\psi(x,\,t) = S(t)\,\psi(x) \qquad (13.3^*)$$

and in the second it will remain

$$\psi(x)$$

To see if L will have a definite value at time $t > 0$ we must check in the first representation whether $\psi(x,t)$ of (13.3*) will be an eigenfunction of L, and in the second whether $\psi(x)$ will be an eigenfunction of $L'(t)$. Hence in the first representation the time dependence is thrown onto the wave functions, and in the second onto the operators. The result is the same whatever representation we use.

To find the time evolution operator $S(t)$ let us adopt the second representation, in which the operators depend on time. Since in this case the time dependence of the state of a system results in a change in the form of the operators, it is natural to interpret the time derivative of an operator for a given quantity as an operator corresponding to the time rate of change of this quantity. Such an interpretation can be taken as a definition for the *time-rate-of-change operator*. We will now find the total time derivative

of $L'(t)$ keeping in mind the possibility of an explicit time dependence of L. We obtain

$$\frac{dL'(t)}{dt} = \dot{S}^+ L S + S^+ \frac{\partial L}{\partial t} S + S^+ L \dot{S} \tag{13.10}$$

where the dot denotes differentiation with respect to time. Now we must return to the first representation. For this we apply to the operator (13.10) a transformation that is the inverse of unitary transformation (13.7) and define dL/dt as

$$\frac{dL}{dt} = S \frac{dL'(t)}{dt} S^+ \tag{13.11}$$

Substituting (13.10) into (13.11), we obtain

$$\frac{dL}{dt} = \frac{\partial L}{\partial t} + S\dot{S}^+ L + L\dot{S}S^+ \tag{13.12}$$

where we used the unitarity of S,

$$SS^+ = 1, \qquad S^+ S = 1 \tag{13.13}$$

Differentiation of the left equation in (13.13) with respect to time yields

$$S\dot{S}^+ + \dot{S}S^+ = 0 \tag{13.14}$$

The last equation shows that the operator $i\dot{S}S^+$, which we will denote as

$$\frac{1}{\hbar} H^* \equiv i\dot{S}S^+ \tag{13.15}$$

will be hermitian. With the help of (13.14) and (13.15), Eq. (13.12) can be translated into

$$\frac{dL}{dt} = \frac{\partial L}{\partial t} + \frac{i}{\hbar}(H^* L - L H^*) \tag{13.16}$$

The second term on the right is the quantum Poisson bracket so that the time derivative of L is

$$\frac{dL}{dt} = \frac{\partial L}{\partial t} + [H^*, L] \tag{13.17}$$

This coincides with the classical expression for the time derivative of L provided that H^* is the classical Hamiltonian function H. We will assume that this is so from the very beginning. Irrespective of the classical analogy this assumption stems from the *law of conservation of energy* and *Bohr's frequency condition*. Indeed, according to the law of conservation of energy we must have

$$\frac{dH}{dt} = \frac{i}{\hbar}(H^* H - H H^*) = 0 \tag{13.18}$$

provided that the Hamiltonian H has no explicit time dependence. This equation will hold for any mechanical system, that is for any form of H, only if

$$H^* = f(H) \qquad (13.19)$$

The form of the function $f(H)$ can be found with the help of Bohr's frequency condition. If we first leave $f(H)$ undefined, we will obtain for the frequency of the light emitted in the transition from the energy level E to the level E' the following expression:

$$v = \frac{1}{2\pi\hbar} [f(E') - f(E)] \qquad (13.20)$$

which coincides with experimental results only if $f(E) = E$.

Conversely, if we were to assume that by classical analogy $H^* \equiv H$, then we could derive the law of conservation of energy and Bohr's frequency condition.

We can thus consider it proved that

$$H = H^* = i\hbar \dot{S} S^+ \qquad (13.21)$$

which transforms the expression (13.16) for the time derivative of operator L into

$$\frac{dL}{dt} = \frac{\partial L}{\partial t} + \frac{i}{\hbar}(HL - LH) \qquad (13.22)$$

or

$$\frac{dL}{dt} = \frac{\partial L}{\partial t} + [H, L] \qquad (13.23)$$

These equations are called the *quantum equations of motion.*

If we assume H to be known, Eqs. (13.3) and (13.21) give the law of the time evolution of the state of a system. Indeed, if we differentiate (13.3) with respect to time, then

$$\frac{\partial \psi(x, t)}{\partial t} = \dot{S}(t) \psi(x, 0)$$

But

$$\psi(x, 0) = S^+(t) \psi(x, t)$$

which yields

$$\frac{\partial \psi(x, t)}{\partial t} = \dot{S} S^+ \psi(x, t) \qquad (13.24)$$

Substituting for $\dot{S}S^+$ its expression (13.21), we find that

$$H\psi - i\hbar \frac{\partial \psi}{\partial t} = 0 \qquad (13.25)$$

This equation came to be known as the *wave equation* although it does not belong to the type of equations that in mathematics are called wave equations.

The wave equation (13.25) can also be obtained by an abstract reasoning. From the point of view of classical mechanics the energy of a system, H, can be considered, up to a difference in sign, to be the generalized momentum that is conjugate to time:

$$H = -p_t \qquad (13.26)$$

If we transfer to quantum mechanics, by analogy with the operators p_x, p_y, p_z we can write

$$p_t = -i\hbar \frac{\partial}{\partial t} \qquad (13.27)$$

By equating the results of applying the operators H and $-p_t$ to a wave function ψ we arrive at Eq. (13.25). We could have said, however, that the form of, say, p_x was found from the condition $[p_x, x] = 1$, whereas for energy and time the Poisson bracket was not considered.

14. Heisenberg's matrices

We can set up the representation in which the time dependence is shifted to the operators (in Section 13 it was the second representation) in the following manner. Let

$$\psi_0(x), \quad \psi_1(x), \ldots, \psi_n(x), \ldots \qquad (14.1)$$

be a complete, orthogonal, and normalized set of functions, for instance, the eigenfunctions of some operator. We wish to find the solution $\psi_n(x, t)$ to the wave equation

$$H\psi - i\hbar \frac{\partial \psi}{\partial t} = 0 \qquad (14.2)$$

the solution satisfying the initial condition

$$\psi_n(x, 0) = \psi_n(x) \qquad (14.3)$$

We can show that for each value of t the solutions

$$\psi_0(x, t), \quad \psi_1(x, t), \ldots, \psi_n(x, t), \ldots \qquad (14.4)$$

form a complete, orthogonal, and normalized set. This is so even if H depends on time explicitly.

We now expand the function $\psi(x, 0)$, which describes the initial state of the system under consideration, in a series involving the functions (14.1):

$$\psi(x, 0) = \sum_{n=0}^{\infty} c_n \psi_n(x) \qquad (14.5)$$

Then at time $t > 0$ the state will be described by

$$\psi(x, t) = \sum_{n=0}^{\infty} c_n \psi_n(x, t) \qquad (14.6)$$

where c_n are the same constants as in (14.5). If we take number n (which labels the function ψ_n) for the independent variable, the state will be described (in this representation) both at $t = 0$ and at $t > 0$ by the same function of n,

$$c(n) = c_n \qquad (14.7)$$

Hence in this representation the time dependence is transferred to the operators, and to find the form of the operators in this representation it is sufficient to shift to the new variable, n.

We can easily find the matrix (kernel) of an operator L in the n-representation. By the general formula (9.13),

$$(n \mid L(t) \mid n') = \int \overline{\psi_n(x, t)} \, L \psi_{n'}(x, t) \, d\tau \qquad (14.8)$$

Such a representation has the property that for the operator dL/dt, which is the *time rate of change* of L, the matrix elements are equal to the time derivative of the corresponding matrix elements of L:

$$\frac{d}{dt}(n \mid L \mid n') = \left(n \left| \frac{dL}{dt} \right| n'\right) \qquad (14.9)$$

This follows from the fact that the time dependence is shifted to the operators. Direct proof of (14.9) is given in Section 4, Chapter IV.

In our discussion we assumed that the ψ_n form a discrete set, the set of eigenfunctions of an operator with a discrete spectrum. But this restriction is not essential. The functions could have been the eigenfunctions of an operator with a continuous spectrum, with n being a continuous parameter. Let us assume, for a moment, that the solution of Eq. (14.2), where we set

$$\psi(x, t) \mid_{t=0} = f(x) \qquad (14.10)$$

can be represented in the form of an integral:

$$\psi(x, t) = \int \psi(x, t; x_0) f(x_0) \, dx_0 \qquad (14.11)$$

Equation (14.11) replaces Eq. (14.6), and $\psi(x, t; x_0)$ takes the part of $\psi_n(x, t)$, parameter x_0 the part of the integer n, and $f(x_0)$ the part of c_n. If we compare the definition (13.3) of the time evolution operator $S(t)$ with (14.11), we see that $\psi(x, t; x_0)$ is the kernel of $S(t)$. We note that in the simpler cases (a free electron, an electron in homogeneous electric field, an oscillator) $\psi(x, t; x_0)$ can be found in closed form. [4]

The representation of operators in which functions (14.1) are the energy eigenfunctions (we assume that the Hamiltonian, H,

---

[4] See, for instance, Chapter XIII in L. de Broglie, *Einführung in die Wellenmechanik*, Akademische Verlagsgesellschaft M. B. H., Leipzig, 1929.

is time independent) plays a special role. This is due to its simplicity and also to its use in studying the radiation of atoms. Let

$$H\psi_n(x) = E_n\psi_n(x) \tag{14.12}$$

The solution of the wave equation (14.2), given the initial conditions (14.3), will obviously be

$$\psi_n(x,\ t) = e^{-iE_n t/\hbar}\psi_n(x) \tag{14.13}$$

If we use these functions to express the matrix elements of a time-varying operator, we find that

$$(n\,|\,L\,(t)\,|\,n') = e^{i\,(E_n-E_{n'})t/\hbar} \int \overline{\psi_n(x)}\,L\psi_{n'}(x)\,d\tau \tag{14.14}$$

The first form of quantum mechanics, discovered by Heisenberg in 1925, was in this "matrix" formulation. Without introducing the operator concept Heisenberg associated with each physical quantity a matrix of type (14.14). We will call these *Heisenberg's matrices*, and the representation in which the time dependence is shifted to the operators corresponding to physical quantities, the *Heisenberg picture*.

Formula (14.13) suggests that a state with a definite energy is a stationary state. Indeed, $\psi_n(x, t)$ of (14.13) remains an eigenfunction of the Hamiltonian for any instant of time, so that if at time zero the energy has a definite value, it will have the same value at subsequent times. This is another way of stating the law of conservation of energy.

We will end this section with a note of historical interest. The first to consider the wave function $\psi(x, y, z, t)$ was Louis de Broglie (1925). De Broglie introduced the idea of the associated wave for an electron and thus founded wave mechanics. Shortly after Schrödinger (1926) in a series of papers offered a mathematical formulation for the problem of stationary atomic states. He showed that the problem reduces to finding the eigenfunctions and eigenvalues of a certain operator (the Hamiltonian). In the same year Schrödinger discovered that his own wave mechanical and Heisenberg's matrix theories are mathematically equivalent. But it was only later that a satisfactory physical interpretation of the wave function was elaborated.

15. Semiclassical approximation

When Planck's constant can be considered small in comparison with quantities of like dimensions encountered in a given problem, one can approximate the solution of the corresponding Schrödinger equation to a solution of the *Hamilton-Jacobi equation* of classical mechanics.

Let us consider the Schrödinger equation

$$-\frac{\hbar^2}{2m}\nabla^2\psi + U(x, y, z)\,\psi = i\hbar\,\frac{\partial\psi}{\partial t} \qquad (15.1)$$

where U is a given function of position, and let us seek its solution in the form

$$\psi = \psi' e^{iS/\hbar} \qquad (15.2)$$

with ψ' a formal power series in $\hbar$. Substituting (15.2) into (15.1) yields

$$\left[\frac{1}{2m}(\text{grad } S)^2 + U + \frac{\partial S}{\partial t}\right]\psi'$$

$$= i\hbar\left[\frac{1}{m}(\text{grad } S \cdot \text{grad } \psi') + \frac{1}{2m}\psi'\nabla^2 S + \frac{\partial\psi'}{\partial t}\right] + \frac{\hbar^2}{2m}\nabla^2\psi' \quad (15.3)$$

If we neglect the term proportional to $\hbar^2$ and set the term not dependent on $\hbar$ equal to zero, we come to two equations

$$\frac{1}{2m}(\text{grad } S)^2 + U + \frac{\partial S}{\partial t} = 0 \qquad (15.4)$$

$$\frac{1}{m}(\text{grad } S \cdot \text{grad } \psi^0) + \frac{1}{2m}\psi^0\,\nabla^2 S + \frac{\partial\psi^0}{\partial t} = 0 \qquad (15.5)$$

In the second for ψ' we substituted its approximated value ψ^0, which corresponds to $\hbar \to 0$.

Equation (15.4) is the Hamilton-Jacobi equation of classical mechanics. Equation (15.5) can be transformed into the *equation of continuity* of classical hydrodynamics. This fact can be proved by multiplying Eq. (15.5) into $2\psi^0$ and assuming that

$$(\psi^0)^2 = \rho \qquad (15.6)$$

We obtain

$$\frac{1}{m}(\text{grad } S \cdot \text{grad}\,\rho) + \frac{\rho}{m}\,\nabla^2 S + \frac{\partial\rho}{\partial t} = 0 \qquad (15.7)$$

or

$$\text{div}\left(\frac{\rho}{m}\,\text{grad } S\right) + \frac{\partial\rho}{\partial t} = 0 \qquad (15.8)$$

In classical mechanics grad $S = \mathbf{p}$ is the momentum vector and $(1/m)\text{grad } S = \mathbf{v}$ is the velocity. Hence Eq. (15.8) can be represented in the following form:

$$\text{div}\,(\rho\mathbf{v}) + \frac{\partial\rho}{\partial t} = 0 \qquad (15.9)$$

which is the equation of continuity.

The solution to the Hamilton-Jacobi equation is commonly called the *action function*, which can be obtained by introducing the *Lagrangian function*

$$\mathscr{L} = \frac{1}{2}m\mathbf{v}^2 - U \qquad (15.10)$$

and evaluating the integral (the *action integral*)

$$S = \int_{t_0}^{t} \mathscr{L}(t)\, dt \qquad (15.11)$$

along the particle's trajectory. To calculate the action integral we first express the Lagrangian function in terms of time and the integration constants of which there are six, because for one particle three second-order Lagrange equations are needed. Integration in (15.11) then yields an action integral S, in terms of time and the initial and final values of coordinates, that satisfies the Hamilton-Jacobi equation.

The solution thus obtained is not unique. Other solutions depend not on the initial values of coordinates but on other constants of integration c_1, c_2, c_3. Furthermore, instead of a rectangular Cartesian coordinate system we can use some other kind of a system, but this ambiguity can be eliminated if we confine ourselves to a rectangular one.

Let

$$S = S(x,\, y,\, z,\, t,\, c_1,\, c_2,\, c_3) \qquad (15.12)$$

be a solution to the Hamilton-Jacobi equation. From classical mechanics we know that

$$\frac{\partial S}{\partial x} = p_x, \qquad \frac{\partial S}{\partial y} = p_y, \qquad \frac{\partial S}{\partial z} = p_z, \qquad \frac{\partial S}{\partial t} = -H \quad (15.13)$$

where p_x, p_y, p_z are the particle's momentum components, and H is the total energy (the Hamiltonian function). Furthermore, the derivatives of S with respect to the constants c_1, c_2, c_3 are new constants, which we denote b_1, b_2, b_3, so that we have

$$\frac{\partial S}{\partial c_1} = b_1, \qquad \frac{\partial S}{\partial c_2} = b_2, \qquad \frac{\partial S}{\partial c_3} = b_3 \qquad (15.14)$$

In the special case when for c_1, c_2, c_3 we take the initial values of coordinates, x_0, y_0, z_0, the constants b_1, b_2, b_3 are the initial values of momentum components but with the opposite sign.

Let us solve Eqs. (15.8) and (15.9) assuming that the solution (15.12) to Eq. (15.4) has already been found. Now we prove that for ρ we can take the determinant

$$\rho = \begin{vmatrix} \dfrac{\partial^2 S}{\partial x\, \partial c_1} & \dfrac{\partial^2 S}{\partial y\, \partial c_1} & \dfrac{\partial^2 S}{\partial z\, \partial c_1} \\[2mm] \dfrac{\partial^2 S}{\partial x\, \partial c_2} & \dfrac{\partial^2 S}{\partial y\, \partial c_2} & \dfrac{\partial^2 S}{\partial z\, \partial c_2} \\[2mm] \dfrac{\partial^2 S}{\partial x\, \partial c_3} & \dfrac{\partial^2 S}{\partial y\, \partial c_3} & \dfrac{\partial^2 S}{\partial z\, \partial c_3} \end{vmatrix} \qquad (15.15)$$

(or, since ρ is positive, its absolute value).

We differentiate (15.4) with respect to the constants included in S, that is, c_1, c_2, c_3, to get

$$\frac{1}{m}\frac{\partial S}{\partial x}\frac{\partial^2 S}{\partial x \partial c_k} + \frac{1}{m}\frac{\partial S}{\partial y}\frac{\partial^2 S}{\partial y \partial c_k} + \frac{1}{m}\frac{\partial S}{\partial z}\frac{\partial^2 S}{\partial z \partial c_k} = -\frac{\partial^2 S}{\partial t \partial c_k} \quad (15.16)$$

with $k = 1, 2, 3$. If we use the fact that

$$\frac{1}{m}\frac{\partial S}{\partial x} = v_x, \qquad \frac{1}{m}\frac{\partial S}{\partial y} = v_y, \qquad \frac{1}{m}\frac{\partial S}{\partial z} = v_z \quad (15.17)$$

we can rewrite Eqs. (15.16) as

$$v_x\frac{\partial^2 S}{\partial x \partial c_k} + v_y\frac{\partial^2 S}{\partial y \partial c_k} + v_z\frac{\partial^2 S}{\partial z \partial c_k} = -\frac{\partial^2 S}{\partial t \partial c_k} \quad (15.18)$$

($k = 1, 2, 3$). The three equations can be solved for the "unknowns" v_x, v_y, v_z, and the determinant of the coefficients of the "unknowns" is ρ [see Eq. (15.5)].

To simplify further formulas let us use the notation (15.14). Equations (15.18) transform into

$$v_x\frac{\partial b_k}{\partial x} + v_y\frac{\partial b_k}{\partial y} + v_z\frac{\partial b_k}{\partial z} = -\frac{\partial b_k}{\partial t} \quad (15.19)$$

(these relations show that the b_k are constant in time, which has been discussed before). The determinant ρ will then be

$$\rho = \begin{vmatrix} \dfrac{\partial b_1}{\partial x} & \dfrac{\partial b_1}{\partial y} & \dfrac{\partial b_1}{\partial z} \\[2mm] \dfrac{\partial b_2}{\partial x} & \dfrac{\partial b_2}{\partial y} & \dfrac{\partial b_2}{\partial z} \\[2mm] \dfrac{\partial b_3}{\partial x} & \dfrac{\partial b_3}{\partial y} & \dfrac{\partial b_3}{\partial z} \end{vmatrix} = \frac{\partial(b_1, b_2, b_3)}{\partial(x, y, z)} \quad (15.20)$$

and ρy_x, ρv_y, ρv_z will then be

$$\rho v_x = -\begin{vmatrix} \dfrac{\partial b_1}{\partial t} & \dfrac{\partial b_1}{\partial y} & \dfrac{\partial b_1}{\partial z} \\[2mm] \dfrac{\partial b_2}{\partial t} & \dfrac{\partial b_2}{\partial y} & \dfrac{\partial b_2}{\partial z} \\[2mm] \dfrac{\partial b_3}{\partial t} & \dfrac{\partial b_3}{\partial y} & \dfrac{\partial b_3}{\partial z} \end{vmatrix} = -\frac{\partial(b_1, b_2, b_3)}{\partial(t, y, z)} \quad (15.21)$$

$$\rho v_y = -\begin{vmatrix} \dfrac{\partial b_1}{\partial x} & \dfrac{\partial b_1}{\partial t} & \dfrac{\partial b_1}{\partial z} \\[2mm] \dfrac{\partial b_2}{\partial x} & \dfrac{\partial b_2}{\partial t} & \dfrac{\partial b_2}{\partial z} \\[2mm] \dfrac{\partial b_3}{\partial x} & \dfrac{\partial b_3}{\partial t} & \dfrac{\partial b_3}{\partial z} \end{vmatrix} = -\frac{\partial(b_1, b_2, b_3)}{\partial(x, t, z)} \quad (15.22)$$

$$\rho v_z = - \begin{vmatrix} \dfrac{\partial b_1}{\partial x} & \dfrac{\partial b_1}{\partial y} & \dfrac{\partial b_1}{\partial t} \\[2mm] \dfrac{\partial b_2}{\partial x} & \dfrac{\partial b_2}{\partial y} & \dfrac{\partial b_2}{\partial t} \\[2mm] \dfrac{\partial b_3}{\partial x} & \dfrac{\partial b_3}{\partial y} & \dfrac{\partial b_3}{\partial t} \end{vmatrix} = - \frac{\partial (b_1, b_2, b_3)}{\partial (x, y, t)} \qquad (15.23)$$

Substituting these into

$$\frac{\partial}{\partial x}(\rho v_x) + \frac{\partial}{\partial y}(\rho v_y) + \frac{\partial}{\partial z}(\rho v_z) + \frac{\partial \rho}{\partial t} \qquad (15.24)$$

we find that (15.24) is identically zero. Therefore the equation of continuity (15.9) is satisfied. Consequently, Eq. (15.5) for the function ψ^0 is also satisfied, with ψ^0 related to ρ by (15.6).

Let us illustrate this theory by the case of a mass point in free motion. Since in free motion velocity is constant and potential energy is zero, we obtain

$$S = \int_0^t \frac{mv^2}{2} dt = \frac{mv^2}{2} t \qquad (15.25)$$

(we have set $t_0 = 0$). For the constants of integration we will take the initial values x_0, y_0, z_0 of the coordinates x, y, z. Then we will have

$$x = x_0 + v_x t, \qquad y = y_0 + v_y t, \qquad z = z_0 + v_z t \qquad (15.26)$$

and hence

$$S = \frac{m}{2t}[(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2] \qquad (15.27)$$

The determinant of the second derivatives of S, which are

$$\frac{\partial^2 S}{\partial x \, \partial x_0} = -\frac{m}{t}, \qquad \frac{\partial^2 S}{\partial y \, \partial y_0} = -\frac{m}{t}, \qquad \frac{\partial^2 S}{\partial z \, \partial z_0} = -\frac{m}{t} \qquad (15.28)$$

(the second derivatives with respect to different coordinates are zero), will be a quantity inversely proportional to t^3. This means that we can set

$$\rho = \frac{\text{constant}}{t^3}, \qquad \rho^{1/2} = \psi^0 = \frac{\text{constant}}{t^{3/2}} \qquad (15.29)$$

which implies that the approximate value of ψ is

$$\psi = \frac{\text{constant}}{t^{3/2}} \exp\left\{ \frac{im}{2\hbar t}[(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2] \right\} \qquad (15.30)$$

Substitution of this expression for ψ into the Schrödinger equation shows that (15.30) is an exact solution and not an approximation. (This can be seen without computation if we use Eq. (15.3) and keep in mind that at $\psi' = \psi^0$, where ψ^0 is given by (15.29), $\nabla^2 \psi' = 0$.)

16. Relation between canonical transformation
and the contact transformation of classical mechanics

For systems that have a classical analog the canonical transformation of operators corresponds to the contact transformation of classical mechanics.

Let q_1, q_2, ..., q_n and p_1, p_2, ..., p_n be the coordinates and momenta prior to transformation and Q_1, Q_2, ..., Q_n and P_1, P_2, ..., P_n the coordinates and momenta after transformation. We consider the case when the transformation function depends on both old and new coordinates:

$$S = S(q_1, ..., q_n; Q_1, ..., Q_n) \qquad (16.1)$$

A *contact transformation* is determined by a relationship between differentials, namely

$$\sum_{r=1}^{n} p_r \, dq_r - \sum_{r=1}^{n} P_r \, dQ_r = dS \qquad (16.2)$$

which implies that

$$p_r = \frac{\partial S}{\partial q_r}, \qquad P_r = - \frac{\partial S}{\partial Q_r} \qquad (16.3)$$

By solving Eqs. (16.3) one can determine q and p in terms of Q and P and the reverse. A solution to Eqs. (16.3) always exists since we assume that the determinant of the derivatives of S does not vanish:

$$\det \left[\frac{\partial^2 S}{\partial q_r \, \partial Q_s} \right] \neq 0 \qquad (16.4)$$

What is the situation in quantum mechanics? The contact transformation of classical mechanics has corresponding to it in quantum mechanics a canonical transformation from a representation in which q is "diagonal" to a representation in which Q is "diagonal". The canonical transformation has the following form. For the sake of brevity [5] we will denote by $\Psi_Q(q)$ the simultaneous eigenfunctions of Q's in terms of q's.

We choose F as the operator we will transform. The kernel or matrix of the transformed operator F^* will then be

$$(Q' \mid F^* \mid Q) = \int \overline{\Psi_{Q'}(q)} \, F \Psi_Q(q) \, dq \qquad (16.5)$$

with dq the product of the differentials:

$$dq = dq_1 \, dq_2 \, ... \, dq_n$$

---

[5] We will often denote the totality of variables $q_1, ..., q_n$ with one symbol q. We use p, Q, P in a similar manner.

The eigenfunction $\Psi_Q(q)$ can be considered the kernel $(q|U|Q)$ of a unitary operator $U = U^{-1}$, and we can write (16.5) as

$$F^* = UFU^{-1} \tag{16.6}$$

Formula (16.5) with $F = 1$ reduces to the orthogonality condition. Then on the left we will find the kernel of the identity operator in terms of Q, namely

$$(Q'|1|Q) = \delta_0(Q - Q') \equiv \delta(Q_1 - Q_1') \ldots \delta(Q_n - Q_n') \tag{16.7}$$

where δ is *Dirac's delta function* [Eq. (16.7) can be considered its definition].

In the semiclassical approximation, for $\Psi_Q(q)$ we can take

$$\Psi_Q(q) = c \left(\left| \frac{\partial^2 S}{\partial q\, \partial Q} \right| \right)^{1/2} e^{iS/\hbar} \tag{16.8}$$

which is a generalization of the expression obtained in the previous section. For the sake of brevity we have introduced the notation

$$\frac{\partial^2 S}{\partial q\, \partial Q} = \det\left[\frac{\partial^2 S}{\partial q_r\, \partial Q_s} \right] \tag{16.9}$$

We also note that in (16.8) the expression in parenthesis is the absolute value of this determinant. The constant c in (16.8) is

$$c = (2\pi\hbar)^{-n/2} \tag{16.10}$$

Let us check to see whether these functions approximately satisfy the orthogonality condition. If we substitute (16.8) into the integral (16.5), then for $F = 1$ the integrand contains a rapidly oscillating factor $e^{i(S - S')/\hbar}$, where S' is obtained from S by substituting Q' for Q. These oscillations cease only if Q' is close to Q. This condition is essential for the integral to be noticeably nonzero. For this reason we can replace $S - S'$ in the exponent by

$$S - S' = - \sum_{r=1}^{n} (Q_r' - Q_r) \frac{\partial S}{\partial Q_r} \tag{16.11}$$

or

$$S - S' = \sum_{r=1}^{n} (Q_r' - Q_r) P_r \tag{16.12}$$

where P_r is defined in (16.3). We can write (16.12) in a short form as

$$S - S' = (Q' - Q) P \tag{16.13}$$

In all the factors of the exponential function we can set $Q' = Q$. We arrive at

$$\int \overline{\Psi_{Q'}(q)}\, \Psi_Q(q)\, dq = c^2 \int e^{i(Q'-Q)P/\hbar} \left| \frac{\partial^2 S}{\partial q\, \partial Q} \right| dq \tag{16.14}$$

But if P_r is defined by (16.3), the determinant in the integrand is the Jacobian of the transformation from variable P to variable q, so that

$$\left| \frac{\partial^2 S}{\partial Q \, \partial q} \right| dq = dP_1 \; \ldots \; dP_n \equiv dP \qquad (16.15)$$

This means that (16.14) can be written in the following form:

$$\int \overline{\Psi_{Q'}(q)} \, \Psi_Q(q) \, dq = c^2 \int e^{i(Q'-Q) \, P/\hbar} \, dP \qquad (16.16)$$

But the right-hand side of (16.16) is simply the product of delta functions (16.7). This finally brings us to

$$\int \overline{\Psi_{Q'}(q)} \, \Psi_Q(q) \, dq = \delta_0(Q - Q') \qquad (16.17)$$

which means that the orthonormality condition is satisfied.

Now let us consider the matrix of an operator F, the operator in terms of q_r and $p_r = -i\hbar(\partial/\partial q_r)$:

$$F = F(q, p) = F\left(q, \; -i\hbar \frac{\partial}{\partial q} \right) \qquad (16.18)$$

When such an operator acts on the exponential function $e^{iS/\hbar}$, the result in the considered approximation will be equal to the product of the exponential function and $F(q, \partial S/\partial q)$:

$$F\left(q, \; -i\hbar \frac{\partial}{\partial q} \right) e^{iS/\hbar} \approx e^{iS/\hbar} F\left(q, \; \frac{\partial S}{\partial q} \right) \qquad (16.19)$$

This also holds for function (16.8). For this reason in (16.5) we can consider F not as a differential operator but as the function in the right-hand side of (16.19). Assuming, as before, that in the factors of the exponential function Q' equals Q, we have

$$(Q'|F^*|Q) = c^2 \int F\left(q, \; \frac{\partial S}{\partial q} \right) e^{i(S-S')/\hbar} \left| \frac{\partial^2 S}{\partial q \, \partial Q} \right| dq \qquad (16.20)$$

We take P for the variables of integration, as in (16.16). Function F will then be transformed thus:

$$F(q, p) = F(q(Q, P), \, p(Q, P)) = F^*(Q, P) \qquad (16.21)$$

where p and P are understood to be the classical expressions (16.3). Because of the approximate nature of formula (16.13) we can write

$$(Q'|F^*|Q) = c^2 \int F^*(Q, \, P) \, e^{i \, (Q'-Q) \, P/\hbar} \, dP \qquad (16.22)$$

To evaluate the integral we first note that multiplying the exponential function in the integrand by P is equivalent to applying operator $-i\hbar(\partial/\partial Q')$ to this function. Hence

$$\int F^*(Q,\ P)\, e^{i\,(Q'-Q)\,P/\hbar}\, dP = \int F^*\Big(Q,\ -i\hbar\,\frac{\partial}{\partial Q'}\Big)\, e^{i\,(Q'-Q)\,P/\hbar}\, dP \qquad (16.23)$$

We can then take F^* outside the integral sign and use (16.16) and (16.17) to get

$$(Q'\,|\,F^*\,|\,Q) = F^*\Big(Q,\ -i\hbar\,\frac{\partial}{\partial Q'}\Big)\,\delta_0\,(Q-Q') \qquad (16.24)$$

Here, as in previous formulas, for the first independent variable of F^* we could have taken Q'. Since the result of applying F^* to a function $\psi(Q)$ is given by the formula

$$F^*\psi\,(Q) = \int (Q\,|\,F^*\,|\,Q')\,\psi\,(Q')\, dQ' \qquad (16.25)$$

we will, by using the expression for the matrix element [Eq. (16.24)], obtain

$$F^*\psi\,(Q) = F^*\Big(Q,\ -i\hbar\,\frac{\partial}{\partial Q}\Big)\,\psi\,(Q) \qquad (16.26)$$

This will be the transformed version of F (to within terms independent of the sequence of factors in F).

We can summarize thus. By using the approximate relation (16.19) we were able to pass from the operator $F(q, -i\hbar(\partial/\partial q))$ to the function $F(q, p)$, which was then expressed in terms of the new variables Q and P via classical formulas for contact transformation. From the new function $F^*(Q, P)$ we returned (when using differentiation with respect to a parameter for calculating the integral) to the operator $F^*(Q, -i\hbar(\partial/\partial Q))$.

We have arrived at the following result. Consider the operator

$$F = F\,(q,\ p) \qquad \text{where } p = -i\hbar\,\frac{\partial}{\partial q} \qquad (16.27)$$

which means that F is expressed in terms of a set of variables q. After a canonical transformation to a new set of variables Q the operator F changes to F^*. By analogy with (16.27) let F^* have the form

$$F^* = F^*\,(Q,\ P) \qquad \text{where } P = -i\hbar\,\frac{\partial}{\partial Q} \qquad (16.28)$$

Now let us assume that the eigenfunctions used to effect the canonical transformation from q to Q in the semiclassical approximation are of form (16.8), so that their phase is $S(q, Q)/\hbar$. In that case F^* can be obtained from F only to within terms inde-

pendent of the sequence of factors [6] by a simple algebraic transformation via

$$F(q, p) = F^*(Q, P) \tag{16.29}$$

$$p = \frac{\partial S}{\partial q}, \qquad P = -\frac{\partial S}{\partial Q} \tag{16.30}$$

with S included in the phase of the unitary transformation. In classical mechanics these formulas represent the contact transformation.

---

[6] The difference of the terms that depend on the sequence of factors will tend to zero as $\hbar \to 0$.

Chapter IV

THE PROBABILISTIC INTERPRETATION
OF QUANTUM MECHANICS

1. Mathematical expectation in the probability theory

We first recall the concept of mathematical expectation known from the probability theory. Let a quantity λ take on the values

$$\lambda_1, \lambda_2, \ldots, \lambda_k, \ldots \tag{1.1}$$

whose probabilities are respectively

$$p_1, p_2, \ldots, p_k, \ldots \tag{1.2}$$

and the sum of the probabilities is unity:

$$p_1 + p_2 + \ldots + p_k + \ldots = 1 \tag{1.3}$$

The *mathematical expectation* of a quantity is the sum of the products of the values of the quantity multiplied by the probabilities of its occurrence:

$$\text{M. E.}\lambda = \sum_k p_k \lambda_k \tag{1.4}$$

where M. E. stands for "mathematical expectation".

Let us use a simple example to illustrate. Suppose we have N lottery tickets. Of these n_1 win λ_1 rubles, n_2 win λ_2 rubles, etc. If this is a lottery in which some holders lose and some win, one of the λ's can be zero. Obviously

$$n_1 + n_2 + \ldots = N \tag{1.5}$$

and if we denote the sum of the winnings by Λ,

$$n_1\lambda_1 + n_2\lambda_2 + \ldots = \Lambda \tag{1.6}$$

The average winning per ticket (if we include nonwinnings) is

$$l = \frac{\Lambda}{N} \tag{1.7}$$

and the probability of winning λ_k rubles is

$$p_k = \frac{n_k}{N} \tag{1.8}$$

If we substitute Λ defined by (1.6) into (1.7) and use (1.8), we can write the expression for the average winning per ticket as

$$l = \sum_k p_k \lambda_k \qquad (1.9)$$

By comparing this with the general formula (1.4) we see that in our example the mathematical expectation is the average winning per ticket:

$$\text{M. E.}\lambda = l \qquad (1.10)$$

In general, the mathematical expectation of a quantity is its *mean value.*

The theory also deals with quantities that vary continuously. Let λ take on, in addition to a discrete set of values, a continuous set in a certain interval. The probability of a value of λ lying between λ and $\lambda + d\lambda$ will, generally speaking, be proportional to $d\lambda$; we set it equal to

$$p(\lambda)\, d\lambda \qquad (1.11)$$

The sum of probabilities must, as before, be equal to unity; this condition can be written as

$$\sum_k p_k + \int p(\lambda)\, d\lambda = 1 \qquad (1.12)$$

where integration is considered over the whole continuous set of values. At last, the mathematical expectation has the form

$$\text{M.E.}\lambda = \sum_k p_k \lambda_k + \int \lambda p(\lambda)\, d\lambda \qquad (1.13)$$

2. Mathematical expectation in quantum mechanics

Let us now turn to the quantum theory. We have already discovered that an electronic state can be described by a wave function ψ. We understood the description to mean that if ψ is an eigenfunction of the operator L corresponding to a physical quantity λ, and λ' is the respective eigenvalue, then specifying ψ is equivalent to indicating that in measuring λ we will get $\lambda = \lambda'$. An eigenvalue is expressed in terms of the corresponding eigenfunction as

$$\lambda' = \frac{\int \bar{\psi} L \psi \, d\tau}{\int \bar{\psi}\psi \, d\tau} \qquad (2.1)$$

So how is one to understand the description of a state by a function ψ that is not, in general, an eigenfunction of a certain

operator L? We will answer this question by introducing a hypothesis about the probabilistic nature of such a description.

Let us assume that we are dealing with a system of electrons all of which are in the same state ψ. If for each electron we measure the quantity λ, then in keeping with our hypothesis separate measurements may give different results (because of the influence of the measuring process on the object). The mean value obtained from these measurements, however, will be a definite number, which will represent the mathematical expectation of λ in state ψ. Hence our hypothesis leads to the assumption that the result of a separate measurement can be accidental but that for a large number of measurements the mean value does not depend on this number if it is great. Since, practically speaking, large numbers of electrons are involved in most cases, the mean value, or the mathematical expectation, of a quantity is even more accessible to measurement than the value of this quantity for a separate electron.

We have thus given a physical interpretation of the concept of mathematical expectation. We will express it in terms of function ψ, which characterizes the electronic state, and operator L, which characterizes the sought-for quantity.

To start with, mathematical expectation must be an invariant, that is, it must not depend on the choice of independent variables in the wave function or on the representation of operators. In short it must be invariant under unitary transformations, considered in the previous chapter.

Aside from this, mathematical expectation must possess two properties known from the theory of probability. First, mathematical expectation of the sum of two quantities must be equal to the sum of the mathematical expectations of these quantities, irrespective of whether the quantities are interdependent or not. Second, if in a given state the quantity λ has a definite value λ' the mathematical expectation must be λ'.

These conditions uniquely determine the expression for mathematical expectation. Invariance implies that mathematical expectation must be expressed in terms of unitary invariants. These are the eigenvalues of operators and expressions of the type

$$\int \bar{\psi} L \psi \, d\tau, \qquad \int \bar{\psi} L^2 \psi \, d\tau, \qquad \text{etc.} \tag{*}$$

But we cannot interpret the eigenvalues of operators as mathematical expectations, if only because an eigenvalue of the sum of two operators is not, generally speaking, equal to the sum of two eigenvalues of the operators. We are thus left with expressions of type (*). From these we must choose the one to the left or a quantity proportional to it, since the first of the properties men-

tioned before requires that the mathematical expectation be linear in L. The second property gives the proportionality factor. Thus [7]

$$\text{M.E.} L = \frac{\int \bar{\psi} L \psi \, d\tau}{\int \bar{\psi} \psi \, d\tau} \qquad (2.2)$$

or, if ψ is normalized,

$$\text{M.E.} L = \int \bar{\psi} L \psi \, d\tau, \qquad \int \bar{\psi} \psi \, d\tau = 1 \qquad (2.3)$$

These expressions satisfy all the stated requirements since, as we know, they are invariant under unitary transformations (the completeness property of eigenfunctions) and also

$$\int \bar{\psi} (L + M) \psi \, d\tau = \int \bar{\psi} L \psi \, d\tau + \int \bar{\psi} M \psi \, d\tau \qquad (2.4)$$

so that

$$\text{M.E.} (L + M) = \text{M.E.} L + \text{M.E.} M \qquad (2.5)$$

Finally, if in state ψ a quantity with an operator L takes on a value λ, that is if

$$L \psi = \lambda \psi \qquad (2.6)$$

then

$$\text{M.E.} L = \lambda \qquad (2.7)$$

Hence this purely mathematical description has led us to a certain expression for the mathematical expectation of a quantity that is defined by an operator L and that characterizes the system in a state ψ. The example of the scattering of α-particles (the Rutherford scattering law), to be considered at the end of Chapter V, Part II, shows that our theory agrees with experiment.

3. The probability formula

Equation (2.2) for mathematical expectation gives us a simple formula for the probability that in measuring a given quantity we will obtain a definite value or a value lying within certain limits.

Let the $\psi(x, \lambda)$ be the set of eigenfunctions of L. We expand $\psi(x)$, which describes the electronic state, in a $\psi(x, \lambda)$-series:

$$\psi(x) = \sum_k c(\lambda_k) \psi(x, \lambda_k) + \int c(\lambda) \psi(x, \lambda) \, d\lambda \qquad (3.1)$$

The result of applying L to ψ is

$$L \psi(x) = \sum_k \lambda_k c(\lambda_k) \psi(x, \lambda_k) + \int \lambda c(\lambda) \psi(x, \lambda) \, d\lambda \qquad (3.2)$$

---

[7] We will denote the quantity by the letter used for its operator.

Assuming that ψ is normalized, we compose the expression for the mathematical expectation of λ. From the completeness condition we have

$$\text{M.E.}\lambda = \int \bar{\psi} L \psi \, d\tau = \sum_k |c(\lambda_k)|^2 \lambda_k + \int \lambda |c(\lambda)|^2 \, d\lambda \qquad (3.3)$$

If we compare this with (1.13), we find that the probability of λ being equal to λ_k is

$$p_k = |c(\lambda_k)|^2 \qquad (3.4)$$

and the probability of λ lying within limits λ and $\lambda + d\lambda$ is

$$p(\lambda) \, d\lambda = |c(\lambda)|^2 \, d\lambda \qquad (3.5)$$

The sum of the probabilities must be unity because due to ψ being normalized and the completeness condition we have

$$\int \bar{\psi} \psi \, d\tau = \sum_k |c(\lambda_k)|^2 + \int |c(\lambda)|^2 \, d\lambda = 1 \qquad (3.6)$$

The expansion coefficients $c(\lambda_k)$ and $c(\lambda)$ are the wave functions describing electronic states in the λ-representation. Thus Eqs. (3.4) and (3.5) give a direct physical interpretation of the squared modulus of the wave function as the probability. For instance, let us set λ to be the coordinates x, y, z. According to (3.5), the electron will be inside a volume with the boundaries

$$(x, x + dx), \qquad (y, y + dy), \qquad (z, z + dz) \qquad (3.7)$$

with a probability

$$|\psi(x, y, z)|^2 \, dx \, dy \, dz \qquad (3.8)$$

Now a general situation with λ a physical quantity. When the initial state is given, to find the probability that the result of measuring λ will be a specific value we must express in the λ variables the wave function that defines this state [to put it differently we must find the coefficients $c(\lambda)$ in the expansion of $\psi(x)$ in a $\psi(x, \lambda)$-series]. The square of its modulus, that is $|c(\lambda)|^2$, gives the probability we are looking for.

Let a state be characterized by a function $\varphi(x, \mu)$ that is an eigenfunction of M corresponding to an eigenvalue μ_k:

$$M\varphi(x, \mu_k) = \mu_k \varphi(x, \mu_k) \qquad (3.9)$$

which means that in this state the measurement of μ gives a definite value, μ_k. What is the probability then that in measuring another quantity, say λ, we will get a value λ_k? If we apply Eq. (3.1) to $\psi(x) = \varphi(x, \mu_k)$ and recall the formula for the expansion coefficients $c(\lambda)$, we find the sought expression

$$|c(\lambda_k)|^2 = \left| \int \overline{\psi(x, \lambda_n)} \varphi(x, \mu_k) \, d\tau \right|^2 \qquad (3.10)$$

On the other hand, if we were to look for the probability that $\mu = \mu_k$ on condition that $\lambda = \lambda_n$, we would come to the same result (3.10). Thus the probability that $\mu = \mu_k$ if $\lambda = \lambda_n$ is equal to the probability that $\lambda = \lambda_n$ if $\mu = \mu_k$.

If λ and μ are the same quantity, then φ and ψ are the eigenfunctions of one operator. According to the orthogonality of eigenfunctions, at $\lambda_k \neq \lambda_n$ the integral will vanish, and at $\lambda_k = \lambda_n$ it will be unity. This agrees with the physical meaning of (3.10) as probability. We see then that two orthogonal wave functions describe incompatible states.

When the given quantity can change continuously, we cannot speak of the probability of it having a definite value — this probability is zero. Instead we can speak of the probability that the quantity lies in a definite interval and also of the "probability density" that is, the probability divided by the width of the interval. For instance,

$$| \psi (x, y, z,) |^2$$

is the probability density for the coordinates.

Here lies the difference between normalization of functions for the discrete and for the continuous spectrum. A transition from eigenfunctions to proper differentials corresponds to the transition from probability density to the probability that the quantity will lie within definite limits.

4. Time dependence of mathematical expectation

The mathematical expectation of a quantity with an operator L,

$$\text{M.E.} L = \frac{\int \bar{\psi} L \psi \, d\tau}{\int \bar{\psi} \psi \, d\tau} \tag{4.1}$$

will, generally speaking, depend on time. If we choose a representation in which the position and momentum operators do not depend on time explicitly, ψ of (4.1) will satisfy the wave equation

$$H\psi - i\hbar \frac{\partial \psi}{\partial t} = 0 \tag{4.2}$$

where H is the Hamiltonian.

Let us first show that the integral in the denominator of (4.1), $\int \bar{\psi} \psi \, d\tau$, will not depend on time. [8] We have

$$\frac{d}{dt} \int \bar{\psi} \psi \, d\tau = \int \frac{\partial \bar{\psi}}{\partial t} \psi \, d\tau + \int \bar{\psi} \frac{\partial \psi}{\partial t} \, d\tau$$

---

[8] Also see (13.6) in Chapter III.

Using the wave equation yields

$$\frac{d}{dt} \int \bar{\psi}\psi \, d\tau = \frac{i}{\hbar} \int (\overline{H\psi}\psi - \bar{\psi}H\psi) \, d\tau$$

But since H is a hermitian operator, the expression on the right is zero. Hence

$$\frac{d}{dt} \int \bar{\psi}\psi \, d\tau = 0 \tag{4.3}$$

If we set

$$\int \bar{\psi}\psi \, d\tau = 1 \tag{4.4}$$

at the initial moment of time ($t = 0$), the normalization (4.4) will remain constant for any $t > 0$. With this in mind we can substitute (4.1) for a simpler formula,

$$\mathbf{M.E.}L = \int \bar{\psi}L\psi \, d\tau \tag{4.5}$$

which also holds for any $t > 0$.

Next we find the time derivative of the mathematical expectation of L. We interchange the derivative and integral signs and for $\partial\bar{\psi}/\partial t$ substitute the wave equation to get

$$\frac{d}{dt} \int \bar{\psi}L\psi \, d\tau = \frac{i}{\hbar} \int \overline{H\psi}L\psi \, d\tau + \int \bar{\psi}\frac{\partial}{\partial t}(L\psi) \, d\tau$$

which in turn yields

$$\frac{d}{dt} \int \bar{\psi}L\psi \, d\tau = \frac{i}{\hbar} \int \bar{\psi}\left(H - i\hbar\frac{\partial}{\partial t}\right)L\psi \, d\tau \tag{4.6}$$

since H is hermitian. If we differentiate explicitly, we get

$$\frac{d}{dt} \int \bar{\psi}L\psi \, d\tau = \int \bar{\psi}\left[\frac{\partial L}{\partial t} + \frac{i}{\hbar}(HL - LH)\right]\psi \, d\tau \tag{4.7}$$

But the expression in brackets is only the operator for the time derivative of L:

$$\frac{dL}{dt} = \frac{\partial L}{\partial t} + \frac{i}{\hbar}(HL - LH) \tag{4.8}$$

Thus Eq. (4.7) expresses the fact that

$$\frac{d}{dt} \int \bar{\psi}L\psi \, d\tau = \int \bar{\psi}\frac{dL}{dt}\psi \, d\tau \tag{4.9}$$

namely, that the time derivative of mathematical expectation equals the mathematical expectation of the time derivative, which is what it should be. If we take this as the starting point, we

come to the formula for the total time derivative of an operator, a result we arrived at earlier in a different way.

In Part II, which is devoted to the Schrödinger equation, we will see that if for L we take operators corresponding to different mechanical quantities, the right-hand side of equations of the type (4.8) will look like the classical expression. Thus in quantum mechanics there exist the same relationships between mathematical expectations and their time derivatives as between quantities and their derivatives in classical mechanics.

5. Correspondence between the theory of linear operators and the quantum theory

Summing up this chapter and Part I, we can say that each quantum mechanical concept has corresponding to it a concept from the theory of linear operators, which implies that we can build a sort of lexicon that will enable us to translate mathematical terms into the language of physics. Roughly it will look like this:

MATHEMATICS	PHYSICS
Linear operator L	Physical quantity λ
Eigenvalues λ' (characteristic, or proper, values)	Observable values of the physical quantity
Eigenfunction ψ (characteristic, or proper, function) for eigenvalue λ'	Mechanical system is in the state with $\lambda = \lambda'$
Commutativity of operators	Possibility to observe physical quantities simultaneously
Squared modulus of eigenfunction, $|\psi|^2$	Probability density
Normalization $\int |\psi|^2 \, d\tau = 1$	Sum of probabilities is unity
Transition to proper differentials for a continuous spectrum	Inequality $\lambda' < \lambda < \lambda' + \Delta\lambda$ has a finite probability
Orthogonality $\int \bar{\varphi}\psi \, d\tau = 0$	States φ and ψ are incompatible
Completeness of the set of functions $\psi(x, \lambda')$	Values λ', λ'', etc. are the only possible ones
Integral $\int \bar{\varphi}L\varphi \, d\tau$	Mathematical expectation of λ in state φ
Squared modulus of coefficient in the expansion of $\varphi(x)$ in a $\varphi(x, \lambda')$-series	Probability that $\lambda = \lambda'$ in state φ

The possibility of such a comparison shows how closely the two theories are related and why the theory of linear operators is so essential to the quantum theory.

6. The concept of statistical ensemble
in quantum mechanics

In the first years of development of quantum mechanics, in the early attempts to find a *statistical (probabilistic) interpretation*, physicists were still bound by the notion of the electron being a classical mass point. Even when de Broglie's idea on the wave nature of matter emerged, waves of matter were at times interpreted as something that carries the mass points. Later, when Heisenberg's relations appeared, they were interpreted as *inaccuracy relations*, and not as uncertainty relations. For instance, it was thought that the electron had definite position and velocity but that there was no possibility of determining either. The square of the modulus of the wave function was interpreted as probability density for a particle — irrespective of the conditions of the actual experiment — to have given coordinates (the coordinates were thought of as definite). A similar interpretation was given to the square of the modulus of the wave function in momentum space. Both probabilities (in ordinary space and in momentum space) were considered simultaneously as the probability of a certain compound event, specifically that the particle has definite values of coordinates and momentum. The actual impossibility, expressed by Heisenberg's relations, of their simultaneous measurement therefore appeared as a paradox or caprice of nature, according to which not everything existing is cognizable.

All these difficulties vanish if we fully admit the dual wave-corpuscular nature of the electron, establish its essence, and grasp what the quantum mechanical probabilities refer to and what statistical ensemble they belong to.

First, let us try to give a general definition of a *statistical ensemble*. We assume an unlimited set of elements having various features, which make it possible to sort these elements and to observe the frequency of occurrence of an element with a given feature. If for this there exists a definite probability (that is, for each element of the set), the set constitutes a statistical ensemble.

In quantum mechanics, as in classical physics, the only sets that can be considered are those whose elements have definite values of the parameters (features) according to which sorting can be done. This implies that the elements of a statistical ensemble must be described in a classical language, and that a quantum object cannot be an element of a statistical ensemble even if a wave function can be ascribed to the object.

The elements of statistical ensembles considered in quantum mechanics are not the micro-objects themselves but the results of experiments with them, a definite experimental arrangement corresponding to a definite ensemble. These results are described

classically and thus can serve as a basis for sorting the elements of the ensemble. Since for different quantities the probability distributions arising from a given wave function correspond to different experimental arrangements, they belong to different ensembles. But the wave function cannot belong to any definite statistical ensemble. This can be illustrated by the following diagram:

	E	p	x	. . .
ψ_1				
ψ_2				
ψ_3				
. . .				

To each square in this diagram belongs a definite statistical ensemble with its own probability distribution for the result of measurement of a given quantity. A row includes the ensembles obtained by measuring various quantities $(E, p, x, \ldots)$, starting from one and the same initial state, whereas a column shows the ensembles obtained by measuring one definite quantity, starting from various states $(\psi_1, \psi_2, \psi_3, \ldots)$.

The deeper reason for the wave function not corresponding to any statistical ensemble is that the concept of the wave function belongs to the potentially possible (that is, to experiments not yet performed, whose outcome and even type are not known). The concept of the statistical ensemble, on the other hand, belongs to the accomplished (to the results of experiments of a definite type already carried out).

The probability of this or that behaviour of an object with a given initial state is determined by the internal properties of the object and by the nature of the external conditions; it is a number

characterizing the potential possibilities of this or that behaviour of the object. And the probability manifests itself in the frequency of occurrence of a given behaviour of the object; the relative frequency is its numerical measure. The probability thus belongs, in essence, to the individual object (and not to an ensemble of objects) and characterizes its potential possibilities. At the same time, to determine its numerical value from experiment one must have the statistics of the realization of these possibilities, so that the experiment must be repeated many times. It is clear from this that the probabilistic character of the quantum theory does not exclude the fact that the theory is based on the properties of an individual object.

To summarize we can say that the purpose of the main concept of quantum mechanics, the concept of a state described by a wave function, is to analyze objectively all potential possibilities inherent in the micro-object. This determines the probabilistic nature of the theory.

Part II

SCHRÖDINGER'S THEORY

Chapter I

THE SCHRÖDINGER EQUATION.
THE HARMONIC OSCILLATOR

1. Equations of motion and the wave equation

As we know, the wave equation, which gives the time dependence of wave functions, must have the following form:

$$H\psi - i\hbar \frac{\partial \psi}{\partial t} = 0 \qquad (1.1)$$

with H the Hamiltonian. Generally speaking, different problems have different Hamiltonians. Schrödinger's theory considers the case when the electron's momentum is small compared to mc, where c is the speed of light. This means that in the absence of a magnetic field the corrections imposed by the theory of relativity can be neglected, so that the electron is moving in an electric field or in a potential $U(x, y, z)$. A generalization of the Schrödinger equation incorporating a magnetic field will be given in Part III.

In Section 8, Chapter III, Part I, by analogy with classical mechanics we wrote the following expression for the Hamiltonian operator:

$$H = \frac{1}{2m} \left(p_x^2 + p_y^2 + p_z^2 \right) + U(x, y, z) \qquad (1.2)$$

where the first term on the right-hand side stands for the kinetic-energy operator, and the second for the potential-energy operator. If for the operators p_x, p_y, p_z we substitute their expressions, we come to the wave equation in the form

$$-\frac{\hbar^2}{2m} \nabla^2 \psi + U(x, y, z) \psi - i\hbar \frac{\partial \psi}{\partial t} = 0 \qquad (1.3)$$

Let us examine the equations of motion that follow from the Schrödinger equation. We find the *velocity* and *acceleration operators*. According to Eq. (13.22), Chapter III, Part I,

$$\frac{dx}{dt} = \frac{i}{\hbar} (Hx - xH)$$

In the expression for H all terms except $p_x^2/(2m)$ commute with x, and therefore

$$\frac{dx}{dt} = \frac{i}{2m\hbar} \left(p_x^2 x - x p_x^2 \right)$$

$$= \frac{i}{2m\hbar} \left[p_x \left(p_x x - x p_x \right) + \left(p_x x - x p_x \right) p_x \right]$$

But we know that

$$[p_x, x] = \frac{i}{\hbar} \left(p_x x - x p_x \right) = 1$$

Hence

$$\frac{dx}{dt} = \frac{1}{m} p_x \tag{1.4}$$

and by analogy

$$\frac{dy}{dt} = \frac{1}{m} p_y, \qquad \frac{dz}{dt} = \frac{1}{m} p_z \tag{1.4*}$$

Thus the velocity operator is the momentum operator divided by m, as we expected. Next we find the operator for dp_x/dt. We have

$$\frac{dp_x}{dt} = \frac{i}{\hbar} \left(H p_x - p_x H \right)$$

The only term in H that does not commute with p_x is $U(x, y, z)$. We then have

$$\frac{dp_x}{dt} \psi = \frac{i}{\hbar} \left(U p_x - p_x U \right) \psi = U \frac{\partial \psi}{\partial x} - \frac{\partial}{\partial x} \left(U \psi \right)$$

$$= - \frac{\partial U}{\partial x} \psi$$

By reasoning in the same way about the other two components in summary we get

$$\frac{dp_x}{dt} = - \frac{\partial U}{\partial x}, \qquad \frac{dp_y}{dt} = - \frac{\partial U}{\partial y}, \qquad \frac{dp_z}{dt} = - \frac{\partial U}{\partial z} \tag{1.5}$$

Equations (1.4) and (1.5) coincide in form with the equations of classical mechanics. If we now recall the interdependence of the equations of motion and the law governing mathematical expectations, we get

$$m \frac{d^2}{dt^2} \int x \bar{\psi} \psi \, d\tau = - \int \frac{\partial U}{\partial x} \bar{\psi} \psi \, d\tau \tag{1.6}$$

and two similar equations for y and z. These equations are called *Ehrenfest's equations*.

2. Constants of the motion

Now let us introduce the concept of the *constants of the motion* into quantum mechanics. It is customary in classical mechanics to call a constant of the motion a quantity (a function of coordinates and momenta) that remains constant under any initial conditions of the problem. In quantum mechanics we can set a constant of the motion to be a physical quantity whose mathematical expectation remains, due to the wave equations, constant under any initial conditions.

For an operator L to be a constant of the motion it is necessary and sufficient that, according to Eqs. (4.7) and (4.8). Chapter IV, Part I, the following conditions hold:

$$\frac{dL}{dt} = \frac{\partial L}{\partial t} + \frac{i}{\hbar}(HL - LH) = 0 \tag{2.1}$$

We can show that if L satisfies (2.1), its eigenfunctions, which are the solutions of the eigenvalue equation

$$L\psi = \lambda\psi \tag{2.2}$$

can be chosen so as at the same time to satisfy the wave equation

$$H\psi - i\hbar\frac{\partial\psi}{\partial t} = 0 \tag{2.3}$$

This will also be the case when the Hamiltonian contains the time variable explicitly. It follows that if at the initial moment ($t = 0$) the physical quantity L had a definite value λ, it will have this value at any subsequent moment.

If operator L does not contain time explicitly, condition (2.1) reduces to the operator's commutativity with the Hamiltonian.

Let us assume, for instance, that $L = H$ and that H does not contain time explicitly. As we know (Section 13, Chapter III, Part I), in this case the energy conservation law holds, that is, a state with a given energy E remains the same for any time t. Equation (2.1) will then be

$$H\psi = E\psi \tag{2.4}$$

The general solution to Eqs. (2.3) and (2.4) will have the following form:

$$\psi = \psi^0(x,\ y,\ z;\ E)\,e^{-iEt/\hbar} \tag{2.5}$$

We can use solutions of type (2.5) to build a solution satisfying the arbitrary initial conditions

$$\psi = f(x,\ y,\ z) \quad \text{at } t = 0 \tag{2.6}$$

For this we must expand the initial wave function (2.6) in a series of the eigenfunctions of the Hamiltonian:

$$f(x, y, z) = \sum_E c(E) \psi^{(0)}(x, y, z; E) \tag{2.7}$$

and then multiply each term by the corresponding exponential factor to get

$$\psi = \sum_E c(E) e^{-iEt/\hbar} \psi^{(0)}(x, y, z; E) \tag{2.8}$$

The series (2.8) will apparently be the general solution of the wave equation satisfying the initial conditions (2.6)[1].

Knowledge of the constants of the motion makes it easier to find the solution to the wave equation. Assume that H does not contain time explicitly and that there are two operators, L and M, that commute with H (which makes them constants of the motion) and, more than that, commute with each other. Then the equations

$$H\psi = E\psi, \qquad L\psi = \lambda\psi, \qquad M\psi = \mu\psi \tag{2.9}$$

will have common eigenfunctions. To find these we can start by solving the simpler of the three equations and then adjust the solution so that it satisfies the other two. We will use this method repeatedly in the future.

3. The Schrödinger equation for the harmonic oscillator

Let us consider the *three-dimensional harmonic oscillator* with the Hamiltonian

$$H = \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2) + \frac{1}{2}m(\omega_1^2 x^2 + \omega_2^2 y^2 + \omega_3^2 z^2) \tag{3.1}$$

A model of this kind can represent a molecule with three vibrational degrees of freedom. To find the eigenfunctions of the Hamiltonian we will use the method mentioned in the previous section and look for operators that will commute with each other and the Hamiltonian. These will obviously be

$$H^{(x)} = \frac{1}{2m} p_x^2 + \frac{1}{2} m\omega_1^2 x^2$$

$$H^{(y)} = \frac{1}{2m} p_y^2 + \frac{1}{2} m\omega_2^2 y^2$$

$$H^{(z)} = \frac{1}{2m} p_z^2 + \frac{1}{2} m\omega_3^2 z^2 \tag{3.2}$$

---

[1] See (14.5) and (14.6), Chapter III, Part I.

7*

To the equation

$$H\psi = E\psi \tag{3.3}$$

we can join the three equations

$$H^{(x)}\psi = E^{(x)}\psi, \qquad H^{(y)}\psi = E^{(y)}\psi, \qquad H^{(z)}\psi = E^{(z)}\psi \tag{3.4}$$

where

$$H = H^{(x)} + H^{(y)} + H^{(z)} \tag{3.5}$$

and

$$E = E^{(x)} + E^{(y)} + E^{(z)} \tag{3.6}$$

Since all three equations in (3.4) are of the same type, it is sufficient to study only one, for instance, the first, which amounts to considering the *one-dimensional oscillator*. Thus for the x-component we have

$$\left(\frac{1}{2m}\, p_x^2 + \frac{1}{2}\, m\omega_1^2 x^2\right)\psi = E^{(x)}\psi \tag{3.7}$$

or

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + \frac{1}{2}\, m\omega_1^2 x^2\psi = E^{(x)}\psi \tag{3.7*}$$

It is convenient to make a change of variables,

$$x\left(\frac{m\omega_1}{\hbar}\right)^{1/2} = \xi, \qquad \frac{E^{(x)}}{\hbar\omega} = \lambda \tag{3.8}$$

Equation (3.7) then reads

$$-\frac{d^2\psi}{d\xi^2} + \xi^2\psi = 2\lambda\psi \tag{3.9}$$

Finally we denote the solution of this equation by $\psi_\lambda^0(\xi)$; the solution to Eq. (3.3) for the three-dimensional oscillator will then be

$$\psi_E(x,\, y,\, z) = \psi_{\lambda_1}^0(\xi_1) \times \psi_{\lambda_2}^0(\xi_2) \times \psi_{\lambda_3}^0(\xi_3) \tag{3.10}$$

where $\xi_1^2 = x^2 m\omega_1/\hbar$, $\xi_2^2 = y^2 m\omega_2/\hbar$, $\xi_3^2 = z^2 m\omega_3/\hbar$, and, according to (3.6),

$$E = \hbar(\omega_1\lambda_1 + \omega_2\lambda_2 + \omega_3\lambda_3) \tag{3.11}$$

Thus the problem reduces itself to an investigation of Eq. (3.9), which we will now undertake.

4. The one-dimensional harmonic oscillator

We start with the equation

$$-\frac{d^2\psi}{d\xi^2} + \xi^2\psi = 2\lambda\psi \tag{4.1}$$

Since the coefficients in (4.1) remain finite for finite ξ, the only singular points of this equation are $\xi = \pm\infty$. We must find solutions that remain finite at $\xi = \pm\infty$. Such solutions, we will see, exist only for specific values of parameter λ; these values are the eigenvalues of the corresponding operator.

To examine the behaviour of Eq. (4.1) as $\xi \to +\infty$ we introduce a new function

$$f = \frac{1}{\psi}\frac{d\psi}{d\xi} \qquad (4.2)$$

which turns (4.1) into

$$\frac{df}{d\xi} + f^2 = \xi^2 - 2\lambda \qquad (4.3)$$

We look for f in the form of a series:

$$f = a\xi + b + \frac{c}{\xi} + \ldots \qquad (4.4)$$

If we now substitute (4.4) into (4.3), we get

$$a^2\xi^2 + 2ab\xi + b^2 + 2ac + a + \ldots = \xi^2 - 2\lambda$$

Identifying the powers of ξ yields

$$a^2 = 1, \qquad b = 0, \qquad (2c + 1)a = -2\lambda \qquad (4.5)$$

The two values of a, namely $a = \pm 1$, give two possible solutions:

$$f = \xi - \frac{\lambda + 1/2}{\xi} + \ldots$$

$$f = -\xi + \frac{\lambda - 1/2}{\xi} + \ldots \qquad (4.4^*)$$

For the first solution, after integrating both sides of (4.8), we get

$$\ln\psi = \frac{1}{2}\xi^2 - \left(\lambda + \frac{1}{2}\right)\ln\xi + \ldots$$

which implies that

$$\psi_1 = e^{\xi^2/2}\xi^{-\lambda-1/2}(1 + \ldots) \qquad (4.6)$$

For the second solution

$$\psi_2 = e^{-\xi^2/2}\xi^{\lambda-1/2}(1 + \ldots) \qquad (4.6^*)$$

where we have left out the terms that decrease as $|\xi|$ increases. For large positive values of ξ the general solution to (4.1) will be

$$\psi = c_1 e^{\xi^2/2}\xi^{-\lambda-1/2}(1 + \ldots) + c_2 e^{-\xi^2/2}\xi^{\lambda-1/2}(1 + \ldots) \qquad (4.7)$$

and for large negative values

$$\psi = c_1' e^{\xi^2/2}\xi^{-\lambda-1/2}(1 + \ldots) + c_2' e^{-\xi^2/2}\xi^{\lambda-1/2}(1 + \ldots) \qquad (4.7^*)$$

What we need is that ψ should remain finite as $\xi \to +\infty$ and $\xi \to -\infty$. This is possible only if

$$c_1 = 0, \qquad c_1' = 0 \tag{4.8}$$

simultaneously.

Hence the solution that interests us must have the following form:

$$\psi = e^{-\xi^2/2} F(\xi) \tag{4.9}$$

where $F(\xi)$ at $\xi = \pm \infty$ must be of the order of $\xi^{\lambda-1/2}$. Let us find the differential equation for $F(\xi)$. Substituting (4.9) into (4.1) and dividing out the exponential factor, we get

$$\frac{d^2F}{d\xi^2} - 2\xi \frac{dF}{d\xi} + (2\lambda - 1) F = 0 \tag{4.10}$$

We look for the solution to (4.10) in the form of a power series in ξ. Since $\xi = 0$ is not a singular point, the series will include only positive powers of ξ, that is

$$F = \sum_{k=0}^{\infty} a_k \xi^k \tag{4.11}$$

which if substituted into (4.10) yields

$$\sum_{k=0}^{\infty} k(k-1) a_k \xi^{k-2} + \sum_{k=0}^{\infty} (-2k + 2\lambda - 1) a_k \xi^k = 0$$

In the first sum the factor $k(k-1)$ vanishes if $k = 0$ or $k = 1$, and we can start the summation process from $k = 2$. If we then change k to $k + 2$, the new k will change from 0 to ∞. The result is

$$\sum_{k=0}^{\infty} [(k+2)(k+1) a_{k+2} + (-2k + 2\lambda - 1) a_k] \xi^k = 0$$

For a power series to vanish all its coefficients must be zero. This gives

$$a_{k+2} = \frac{2k - 2\lambda + 1}{(k+2)(k+1)} a_k \tag{4.12}$$

This formula gives us a rule for successive determination of the a_k, with the first two, a_0 and a_1, being arbitrary constants. The series for $F(\xi)$ will then be

$$F(\xi) = a_0 \left(1 + \frac{1-2\lambda}{1 \times 2} \xi^2 + \frac{(1-2\lambda)(5-2\lambda)}{1 \times 2 \times 3 \times 4} \xi^4 + \cdots \right)$$

$$+ a_1 \left(\xi + \frac{3-2\lambda}{2 \times 3} \xi^3 + \frac{(3-2\lambda)(7-2\lambda)}{2 \times 3 \times 4 \times 5} \xi^5 + \cdots \right) \tag{4.13}$$

or

$$F(\xi) = a_0 F_0(\xi) + a_1 F_1(\xi) \tag{4.14}$$

where F_0 and F_1 denote the corresponding series. As $|\xi| \to \infty$ the function $F(\xi)$ must be of the order of $|\xi|^{\lambda-\frac{1}{2}}$ for both positive and negative values of ξ. But

$$F(-\xi) = a_0 F_0(\xi) - a_1 F_1(\xi) \tag{4.15}$$

Hence $a_0 F_0(\xi)$ and $a_1 F_1(\xi)$ must each be of an order no higher than $\xi^{\lambda-\frac{1}{2}}$. Here two cases are possible: either F_0 and F_1 are infinite series or one at least is terminated. In the first, both series converge for all ξ, since the ratio of two successive terms

$$\frac{a_{k+2}\xi^{k+2}}{a_k\xi^k} = \frac{2k - 2\lambda + 1}{(k+2)(k+1)} \xi^2 \tag{4.16}$$

tends to zero as $k \to \infty$. But the same ratio shows that starting from a definite term (for which $k > \lambda - \frac{1}{2}$) all terms will have the same sign. Hence the expansion will contain terms of the same sign with arbitrarily large powers of ξ, and its sum will grow faster than any finite power of ξ which contradicts our assumption (the order must not be greater than $\xi^{\lambda-\frac{1}{2}}$). It follows that at least one of the series, F_0 or F_1, must terminate. This means that for a given k, for instance, $k = n$, the expansion coefficient a_{k+2} vanishes but a_k is not zero. By the recursion relation (4.12) this will happen if

$$\lambda = n + \frac{1}{2}, \qquad n = 0, 1, 2, \ldots \tag{4.17}$$

If n is even, $F_0(\xi)$ becomes a polynomial. If we put $a_1 = 0$, we get a solution that satisfies our assumption. On the other hand, for n odd it is $F_1(\xi)$ that becomes a polynomial, and we must set $a_0 = 0$. In both cases the solution will be a polynomial of degree n.

5. Hermite polynomials

Polynomials that satisfy the equation

$$\frac{d^2 F}{d\xi^2} - 2\xi \frac{dF}{d\xi} + 2nF = 0 \tag{5.1}$$

with integral n are called the *Hermite polynomials* and are denoted $H_n(\xi)$. The function (4.13) gives for them a representation in the form of a series; namely, for n even

$$H_n(\xi) = a_0 \left(1 - \frac{2n}{1 \times 2} \xi^2 + \frac{2n(2n-4)}{1 \times 2 \times 3 \times 4} \xi^4 - \ldots \right) \tag{5.2}$$

and for n odd

$$H_n(\xi) = a_1 \left(\xi - \frac{2n-2}{2 \times 3} \xi^3 + \frac{(2n-2)(2n-6)}{2 \times 3 \times 4 \times 5} \xi^5 - \ldots \right) \tag{5.3}$$

The constants a_0 and a_1 are usually defined in a way such that the leading coefficient is 2^n. This implies that

$$a_0 = (-1)^{n/2} \frac{n!}{(n/2)!}, \qquad n \text{ even} \tag{5.4}$$

$$a_1 = 2(-1)^{(n-1)/2} \frac{n!}{[(n-1)/2]!}, \qquad n \text{ odd} \tag{5.5}$$

If we rearrange the polynomials $H_n(\xi)$ according to decreasing powers of ξ, we get an expression valid for both n even and n odd:

$$H_n(\xi) = (2\xi)^n - \frac{n(n-1)}{1}(2\xi)^{n-2}$$
$$+ \frac{n(n-1)(n-2)(n-3)}{1 \times 2}(2\xi)^{n-4} - \ldots \tag{5.6}$$

Let us show that Hermite polynomials can be represented as

$$H_n(\xi) = (-1)^n e^{\xi^2} \frac{d^n}{d\xi^n} e^{-\xi^2} \tag{5.7}$$

First, it is easy to see that the right-hand side is a polynomial whose leading term is $(2\xi)^n$. Indeed, this term originates from the nth power of the derivative of the exponent $-\xi^2$. Since Eq. (5.1) has only one solution in the form of a polynomial, what remains to be shown is that the right-hand side of (5.7) satisfies (5.1). For this we note that $y = e^{-\xi^2}$ is the solution to

$$y' + 2\xi y = 0$$

Differentiating this equation $n+2$ times yields

$$y^{(n+2)} + 2\xi y^{(n+1)} + (2n+2) y^{(n)} = 0$$

or

$$z'' + 2\xi z' + (2n+2) z = 0$$

where

$$z = y^{(n)}$$

If we then introduce a new function

$$w = e^{\xi^2} z = e^{\xi^2} \frac{d^n}{d\xi^n} e^{-\xi^2}$$

we find that w is the solution to the equation

$$w'' - 2\xi w' + 2nw = 0$$

which coincides with Eq. (5.1) for $H_n(\xi)$. Hence the representation (5.7) is valid.

Next we differentiate this equation, namely

$$\frac{d^2 H_n}{d\xi^2} - 2\xi \frac{d H_n}{d\xi} + 2n H_n = 0 \tag{5.8}$$

with respect to ξ and get

$$\frac{d^2 H'_n}{d\xi^2} - 2\xi \frac{dH'_n}{d\xi} + (2n - 2) H'_n = 0$$

We have just proved that $H'_n(\xi)$ and $H_{n-1}(\xi)$ satisfy the same equation and differ only by a factor. Since the leading term in $H'_n(\xi)$ is $2n(2\xi)^{n-1}$, and in $H_{n-1}(\xi)$ it is $(2\xi)^{n-1}$, we have the following relation:

$$\frac{dH_n}{d\xi} = 2n H_{n-1} \tag{5.9}$$

On the other hand, if we differentiate (5.7), we have

$$\frac{dH_n}{d\xi} = 2\xi H_n - H_{n+1} \tag{5.10}$$

Comparing the two expressions for the derivative, we come to a recursion relation that connects three successive Hermite polynomials:

$$H_{n+1} - 2\xi H_n + 2n H_{n-1} = 0 \tag{5.11}$$

The functions

$$\psi_n(\xi) = c_n e^{-\xi^2/2} H_n(\xi) \tag{5.12}$$

are often referred to as the *Hermite functions* and are the eigenfunctions of the operator in the left-hand side of

$$-\frac{d^2\psi_n}{d\xi^2} + \xi^2 \psi_n = (2n + 1) \psi_n \tag{5.13}$$

and thus are orthogonal:

$$\int_{-\infty}^{+\infty} \psi_n(\xi) \psi_{n'}(\xi) \, d\xi = 0 \quad \text{at} \quad n \neq n' \tag{5.14}$$

For them to be normalized we must determine the constant c_n from the condition

$$\int_{-\infty}^{+\infty} \psi_n^2(\xi) \, d\xi = 1 \tag{5.15}$$

or

$$\int_{-\infty}^{+\infty} e^{-\xi^2} H_n^2(\xi) \, d\xi = \frac{1}{c_n^2} \tag{5.16}$$

Let us evaluate the last integral. Substituting (5.7) for $H_n(\xi)$, we have

$$\frac{1}{c_n^2} = (-1)^n \int_{-\infty}^{+\infty} \frac{d^n}{d\xi^n} (e^{-\xi^2}) H_n(\xi) \, d\xi$$

Atter we integrate n times by parts, we get

$$\frac{1}{c_n^2} = \int_{-\infty}^{+\infty} e^{-\xi^2} \frac{d^n H_n}{d\xi^n} \, d\xi$$

But $d^n H_n / d\xi^n = 2^n n!$, which implies that

$$\frac{1}{c_n^2} = 2^n n! \int_{-\infty}^{+\infty} e^{-\xi^2} \, d\xi = \pi^{1/2} 2^n n! \qquad (5.17)$$

Consequently the functions

$$\psi_n(\xi) = \frac{1}{\pi^{1/4} (2^n n!)^{1/2}} e^{-\xi^2/2} H_n(\xi) \qquad (5.18)$$

are orthonormalized. If we then solve (5.18) for $H_n(\xi)$ and substitute the result into (5.9) and (5.11), we get

$$\frac{d\psi_n}{d\xi} + \xi\psi_n = (2n)^{1/2} \psi_{n-1} \qquad (5.19)$$

$$\xi\psi_n = \left(\frac{n}{2}\right)^{1/2} \psi_{n-1} + \left(\frac{n+1}{2}\right)^{1/2} \psi_{n+1} \qquad (5.20)$$

Then we see that substituting (5.20) into (5.19) yields

$$\frac{d\psi_n}{d\xi} = \left(\frac{n}{2}\right)^{1/2} \psi_{n-1} - \left(\frac{n+1}{2}\right)^{1/2} \psi_{n+1} \qquad (5.21)$$

In conclusion we will give without proof the asymptotic form of $\psi_n(\xi)$ for $2n - \xi^2 \gg 1$:

$$\psi_n(\xi) = \frac{(2/\pi)^{1/2}}{(2n - \xi^2)^{1/4}} \cos\left[\left(n + \frac{1}{2}\right) \arcsin \frac{\xi}{(2n)^{1/2}}\right.$$
$$\left. + \frac{\xi(2n - \xi^2)^{1/2}}{2} - \frac{n\pi}{2}\right] \qquad (5.22)$$

6. Canonical transformation as illustrated by the harmonic-oscillator problem

When solving the oscillator problem, we took ξ for the independent variable, thus expressing the state of the oscillator by a function of ξ, namely $\psi(\xi)$. Now let us take for the independent variable the quantum number n, which labels the oscillator's energy levels. We expand $\psi(\xi)$ in a (complete) set of the energy eigenfunctions:

$$\psi(\xi) = \sum_{n=0}^{\infty} c_n \psi_n(\xi) \qquad (6.1)$$

where the expansion coefficients are determined by simple integration:

$$c_n = \int_{-\infty}^{+\infty} \overline{\psi_n(\xi)}\, \psi_n(\xi)\, d\xi \qquad (6.2)$$

These coefficients, we know, can be interpreted as the wave function in terms of n.

The next step is to find the form of our two simplest operators

$$\xi = \left(\frac{m\omega}{\hbar}\right)^{1/2} x, \qquad p_\xi = -i\frac{\partial}{\partial\xi} = \frac{1}{(m\omega\hbar)^{1/2}} p_x \qquad (6.3)$$

in term of the new variable. We have

$$\xi\psi(\xi) = \sum_{n=0}^{\infty} c_n \xi\psi_n(\xi)$$

If we use (5.20) for $\xi\psi_n$ and collect like terms, we get

$$\xi\psi(\xi) = \sum_{n=0}^{\infty}\left[\left(\frac{n}{2}\right)^{1/2} c_{n-1} + \left(\frac{n+1}{2}\right)^{1/2} c_{n+1}\right]\psi_n(\xi) \qquad (6.4)$$

This means that the operator ξ transforms a function with expansion coefficients c_n into a function with expansion coefficients c'_n, where

$$\xi c_n = c'_n = \left(\frac{n}{2}\right)^{1/2} c_{n-1} + \left(\frac{n+1}{2}\right)^{1/2} c_{n+1} \qquad (6.5)$$

(the symbol ξ is understood to be an operator). The last relationship can be written as

$$c'_n = \sum_k (n\,|\,\xi\,|\,k)\, c_k \qquad (6.6)$$

where

$$(n\,|\,\xi\,|\,k) = \left(\frac{n}{2}\right)^{1/2} \delta_{n-1\,k} + \left(\frac{n+1}{2}\right)^{1/2} \delta_{n+1\,k} \qquad (6.7)$$

which implies that

$$(n\,|\,\xi\,|\,n-1) = \left(\frac{n}{2}\right)^{1/2}, \qquad (n\,|\,\xi\,|\,n+1) = \left(\frac{n+1}{2}\right)^{1/2} \qquad (6.8)$$

while all other terms in (6.6) vanish. Thus the position operator, ξ, can be expressed as a matrix with elements (6.7). Written out in full

$$\xi = \begin{pmatrix} 0 & \sqrt{1/2} & 0 & 0 & 0 & \cdots \\ \sqrt{1/2} & 0 & \sqrt{1} & 0 & 0 & \cdots \\ 0 & \sqrt{1} & 0 & \sqrt{3/2} & 0 & \cdots \\ 0 & 0 & \sqrt{3/2} & 0 & \sqrt{2} & \cdots \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix} \qquad (6.9)$$

We turn to the operator

$$p_\xi = -i\frac{d}{d\xi}$$

Reasoning along the same lines, we come to

$$p_\xi \psi = -i\sum_{n=0}^{\infty} c_n \frac{d\psi_n}{d\xi}$$

then use (5.21) for $d\psi_n/d\xi$ and collect like terms. The result is

$$p_\xi \psi = \sum_{n=0}^{\infty} \left[i\left(\frac{n}{2}\right)^{1/2} c_{n-1} - i\left(\frac{n+1}{2}\right)^{1/2} c_{n+1}\right]\psi_n \qquad (6.10)$$

Thus the operator p_ξ can be expressed as a matrix with elements

$$(n\,|\,p_\xi\,|\,k) = i\left(\frac{n}{2}\right)^{1/2}\delta_{n-1\,k} - i\left(\frac{n+1}{2}\right)^{1/2}\delta_{n+1\,k} \qquad (6.11)$$

In full form this matrix is

$$p_\xi = \begin{bmatrix} 0 & -i\sqrt{1/2} & 0 & \cdots \\ i\sqrt{1/2} & 0 & -i\sqrt{1/2} & \cdots \\ 0 & i\sqrt{1} & 0 & \cdots \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \end{bmatrix} \qquad (6.12)$$

The Hamiltonian (in dimensionless units)

$$H = \frac{1}{2}(p_\xi^2 + \xi^2) \qquad (6.13)$$

expressed in terms of the variable, must reduce to a multiplication operator. To verify this we first note that

$$H\psi = \sum_{n=0}^{\infty} c_n\left(-\frac{1}{2}\frac{d^2\psi_n}{d\xi^2} + \frac{1}{2}\xi^2\psi_n\right) = \sum_{n=0}^{\infty}\left(n+\frac{1}{2}\right)c_n\psi_n$$

in virtue of the differential equation (5.13) for ψ_n. It follows then that

$$Hc_n = \left(n+\frac{1}{2}\right)c_n \qquad (6.14)$$

The matrix elements of H, which is obviously a diagonal matrix, are

$$(n\,|\,H\,|\,k) = \left(n+\frac{1}{2}\right)\delta_{nk} \qquad (6.15)$$

We can obtain the same results if we square matrices ξ and p_ξ. We have

$$(n\,|\,\xi^2\,|\,n') = \sum_k (n\,|\,\xi\,|\,k)\,(k\,|\,\xi\,|\,n')$$

$$= \sum_k \left[\left(\frac{n}{2}\right)^{1/2}\delta_{n-1\,k} + \left(\frac{n+1}{2}\right)^{1/2}\delta_{n+1\,k}\right]$$

$$\times \left[\left(\frac{k}{2}\right)^{1/2}\delta_{k-1\,n'} + \left(\frac{k+1}{2}\right)^{1/2}\delta_{k+1\,n'}\right]$$

$$= \frac{1}{2}\,[n\,(n-1)]^{1/2}\,\delta_{n-2\,n'} + \left(n+\frac{1}{2}\right)\delta_{nn'}$$

$$+ \frac{1}{2}\,[(n+1)\,(n+2)]^{1/2}\,\delta_{n+2\,n'} \qquad (6.16)$$

and similarly

$$(n\,|\,p_\xi^2\,|\,n') = -\sum_k\left[\left(\frac{n}{2}\right)^{1/2}\delta_{n-1\,k} - \left(\frac{n+1}{2}\right)^{1/2}\delta_{n+1\,k}\right]$$

$$\times\left[\left(\frac{k}{2}\right)^{1/2}\delta_{k-1\,n'} - \left(\frac{k+1}{2}\right)^{1/2}\delta_{k+1\,n'}\right]$$

$$= -\frac{1}{2}\,[n\,(n-1)]^{1/2}\,\delta_{n-2\,n'} + \left(n+\frac{1}{2}\right)\delta_{nn'}$$

$$- \frac{1}{2}\,[(n+1)\,(n+2)]^{1/2}\,\delta_{n+2\,n'} \qquad (6.17)$$

The half-sum of (6.16) and (6.17) is

$$\frac{1}{2}\,[(n\,|\,\xi^2\,|\,n') + (n\,|\,p_\xi^2\,|\,n')] = (n\,|\,H\,|\,n') = \left(n+\frac{1}{2}\right)\delta_{nn'} \qquad (6.15^*)$$

which is what it should be.

The commutation relation for ξ and p_ξ,

$$i\,(p_\xi\xi - \xi p_\xi) = 1 \qquad (6.18)$$

must also hold for our matrices. Indeed, according to the rule of matrix multiplication,

$$(n\,|\,p_\xi\xi\,|\,n') = \frac{i}{2}\,[n\,(n-1)]^{1/2}\,\delta_{n-2\,n'} - \frac{i}{2}\,\delta_{nn'}$$

$$- \frac{i}{2}\,[(n+1)\,(n+2)]^{1/2}\,\delta_{n+2\,n'}$$

$$(n\,|\,\xi p_\xi\,|\,n') = \frac{i}{2}\,[n\,(n-1)]^{1/2}\,\delta_{n-2\,n'} + \frac{i}{2}\,\delta_{nn'}$$

$$- \frac{i}{2}\,[(n+1)\,(n+2)]^{1/2}\,\delta_{n+2\,n'}$$

Hence

$$(n\,|\,i\,(p_\xi\xi - \xi p_\xi)\,|\,n') = \delta_{nn'} \qquad (6.19)$$

which was what we set out to prove.

7. Heisenberg's uncertainty relations

The mathematical expectation of a certain physical quantity L that is in the nth state of an oscillator is

$$\text{M. E. } L = \int \overline{\psi_n} L \psi_n \, d\xi \tag{7.1}$$

This is obviously the diagonal element $(n|L|n)$ of the matrix that corresponds to the operator L in the n-representation. It follows then that the mathematical expectations of the coordinate $x = \xi(\hbar/m\omega)^{1/2}$ and the momentum $p_x = p_\xi(m\hbar\omega)^{1/2}$ are zero. Indeed, the respective diagonal elements vanish:

$$\overline{x} = 0, \qquad \overline{p}_x = 0 \tag{7.2}$$

Let us now find the mathematical expectations of the squares of the deviations of x and p_x from the means $\overline{x}$ and $\overline{p}_x$. We have

$$\text{M. E. } (x - \overline{x})^2 = \text{M. E. } x^2 = \frac{\hbar}{m\omega} \text{M. E. } \xi^2$$

$$= \frac{\hbar}{m\omega} (n|\xi^2|n) \tag{7.3}$$

so that, if we denote the left-hand side as $(\Delta x)^2$, we get

$$(\Delta x)^2 = \text{M. E. } (x - \overline{x})^2 = \frac{\hbar}{m\omega} \left(n + \frac{1}{2}\right) \tag{7.4}$$

Similarly

$$\text{M. E. } (p_x - \overline{p}_x)^2 = \text{M. E. } p_x^2 = m\hbar\omega (n|p_\xi^2|n) \tag{7.5}$$

or

$$(\Delta p_x)^2 = m\hbar\omega \left(n + \frac{1}{2}\right) \tag{7.6}$$

Thus

$$\Delta x = \left[\frac{\hbar}{m\omega} \left(n + \frac{1}{2}\right)\right]^{1/2} \tag{7.7}$$

$$\Delta p_x = \left[m\hbar\omega \left(n + \frac{1}{2}\right)\right]^{1/2} \tag{7.8}$$

and we come to the final result

$$\Delta p_x \Delta x = \left(n + \frac{1}{2}\right) \hbar \tag{7.9}$$

We can interpret the two quantities, Δp_x and Δx, as the *root-mean-square*, or *standard*, *deviations* of the measured values of the oscillator's momentum and position from the mathematical expectations of these quantities.

What we have proved is a very general result. If we understand (7.9) as a relationship between the orders of magnitude of the standard deviations and if we introduce a quantum number n in

the proper way, the formula will be valid not only for an oscillator but for any system in its nth state. The product of the standard deviations will be minimal in the ground state ($n = 0$), which implies that the following inequality also holds:

$$\Delta p_x \, \Delta x \geqslant \hbar/2 \tag{7.10}$$

We will now prove that this inequality holds not only for an electron in one of the oscillator's eigenstates but for an electron in any state. To this end we will take the mathematical expectations of x and p_x equal to zero (this restriction can be easily lifted).

Let the electron be in a certain state $\psi(x)$. We introduce two real constants α and β and consider the inequality

$$\int\limits_{-\infty}^{+\infty} \left| \alpha x \psi + \beta \frac{d\psi}{dx} \right|^2 dx \geqslant 0 \tag{7.11}$$

which holds for all values of α and β. Evaluating the square of the modulus in the integrand, we come to

$$\alpha^2 \int x^2 \bar{\psi} \psi \, dx + \alpha\beta \int x \left(\frac{d\bar{\psi}}{dx} \psi + \bar{\psi} \frac{d\psi}{dx} \right) dx + \beta^2 \int \frac{d\bar{\psi}}{dx} \frac{d\psi}{dx} \, dx \geqslant 0 \tag{7.12}$$

or

$$A\alpha^2 - B\alpha\beta + C\beta^2 \geqslant 0 \tag{7.13}$$

where

$$A = \int x^2 \bar{\psi} \psi \, dx$$

$$B = - \int x \frac{d}{dx} (\bar{\psi}\psi) \, dx = \int \bar{\psi}\psi \, dx$$

$$C = \int \frac{d\bar{\psi}}{dx} \frac{d\psi}{dx} \, dx = - \int \bar{\psi} \frac{d^2\psi}{dx^2} \, dx \tag{7.14}$$

For the quadratic form (7.13) to be positive it is necessary that $4AC \geqslant B^2$, or

$$A^{1/2} C^{1/2} \geqslant B/2 \tag{7.15}$$

since A, B and C are all positive. But according to (7.14) we have

$$A^{1/2} = \Delta x, \qquad C^{1/2} = \Delta p_x/\hbar, \qquad B = 1 \tag{7.16}$$

whence

$$\Delta p_x \, \Delta x \geqslant \hbar/2$$

which is what we set out to prove.

Similar inequalities evidently hold for the other two coordinates, so that in summary we have

$$\Delta p_x \, \Delta x \geqslant \hbar/2, \qquad \Delta p_y \, \Delta y \geqslant \hbar/2, \qquad \Delta p_z \, \Delta z \geqslant \hbar/2 \tag{7.17}$$

The inequalities (7.17) were first stated by Heisenberg, who in a number of examples from physics showed how an increase in the precision of measuring the position leads to a decrease in the precision of measuring the momentum, and vice versa. We note that the above formal derivation of (7.17) belongs to Hermann Weyl.

8. The time dependence of matrices. A comparison with the classical theory

The canonical transformation of Section 6 did not contain time and thus gave a representation in which the mathematical form of operators does not depend on time (see Section 13, Chapter III, Part I). We now turn to a representation in which the time dependence is, so to say, shifted to the operators themselves. For this we must find a unitary operator $S(t)$ such that will transform the initial ($t = 0$) state ψ into a state at time $t > 0$:

$$\psi(x, t) = S(t)\, \psi(x,\ 0) \tag{8.1}$$

Then an operator L as a function of time will be represented as

$$L'(t) = S^+(t)\, L S(t) \tag{8.2}$$

If for the independent variable we take the energy, $S(t)$ assumes the simplest form. In this variable the wave equation

$$H\psi - i\hbar\, \frac{\partial \psi}{\partial t} = 0$$

reads

$$Hc_n - i\hbar \frac{\partial c_n}{\partial t} = 0 \tag{8.3}$$

or

$$\hbar\omega \left(n + \frac{1}{2}\right) c_n - i\hbar\, \frac{\partial c_n}{\partial t} = 0 \tag{8.4}$$

(Here we have shifted from the dimensionless units of Section 6 to conventional units.) The solution to this equation is

$$c_n = c_n^0 \exp\left[-i\,(n + {}^1\!/_2)\, \omega t\right] \tag{8.5}$$

Thus the application of $S(t)$ reduces to multiplying c_n^0 by the exponential $\exp\left[-i(n + {}^1\!/_2)\omega t\right]$. This implies that $S(t)$ can be represented by a diagonal matrix

$$S(t) = \begin{Bmatrix} e^{-i\omega t/2} & 0 & 0 \ldots \\ 0 & e^{-3i\omega t/2} & 0 \ldots \\ 0 & 0 & e^{-5i\omega t/2} \\ \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \end{Bmatrix} \tag{8.6}$$

with matrix elements

$$(n \mid S(t) \mid n') = \exp\left[-i(n + \tfrac{1}{2})\,\omega t\right]\delta_{nn'} \qquad (8.7)$$

Using (8.2) to calculate Heisenberg's matrix $x(t)$, we get

$(n \mid x(t) \mid n')$

$$= e^{i(n+1/2)\omega t}\left[\left(\frac{n\hbar}{2m\omega}\right)^{1/2}\delta_{n-1\,n'} + \left(\frac{(n+1)\hbar}{2m\omega}\right)^{1/2}\delta_{n+1\,n'}\right]e^{-i(n'+1/2)\omega t}$$

or

$(n \mid x(t) \mid n')$

$$= e^{i\omega t}\left(\frac{n\hbar}{2m\omega}\right)^{1/2}\delta_{n-1\,n'} + e^{-i\omega t}\left(\frac{(n+1)\hbar}{2m\omega}\right)^{1/2}\delta_{n+1\,n'} \qquad (8.8)$$

which leads to the following matrix:

$$x(t) = \begin{vmatrix} 0 & \left(\frac{\hbar}{2m\omega}\right)^{1/2}e^{-i\omega t} & 0 & \cdots \\ \left(\frac{\hbar}{2m\omega}\right)^{1/2}e^{i\omega t} & 0 & \left(\frac{\hbar}{m\omega}\right)^{1/2}e^{-i\omega t} & \cdots \\ 0 & \left(\frac{\hbar}{m\omega}\right)^{1/2}e^{i\omega t} & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{vmatrix} \qquad (8.9)$$

In the same fashion

$$(n \mid p_x(t) \mid n') = ie^{i\omega t}\left(\frac{n\hbar m\omega}{2}\right)^{1/2}\delta_{n-1\,n'} - ie^{-i\omega t}\left(\frac{(n+1)\hbar m\omega}{2}\right)^{1/2}\delta_{n+1\,n'} \qquad (8.10)$$

and

$$p_x(t) = \begin{vmatrix} 0 & -i\left(\frac{\hbar m\omega}{2}\right)^{1/2}e^{i\omega t} & 0 & \cdots \\ i\left(\frac{\hbar m\omega}{2}\right)^{1/2}e^{i\omega t} & 0 & -i(\hbar m\omega)^{1/2}e^{i\omega t} & \cdots \\ 0 & i(\hbar m\omega)^{1/2}e^{i\omega t} & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{vmatrix} \qquad (8.11)$$

If according to (8.2) we were to form the operator

$$H'(t) = S^+(t)\,HS(t)$$

we would discover that its matrix coincides with that for H and hence does not depend on time. This is a logical result, since the oscillator's energy remains constant.

It can easily be shown that Heisenberg's matrices satisfy the following equations of motion:

$$\frac{dx}{dt} = \frac{1}{m}\,p_x, \qquad \frac{dp_x}{dt} = -m\omega^2 x \qquad (8.12)$$

where dx/dt and dp_x/dt are matrices whose elements are derivatives of the matrix elements of $x(t)$ and $p_x(t)$. Equations (8.12) coincide in form with the classical equations.

The matrix elements of $x(t)$ and $p_x(t)$ resemble the terms in the *Fourier series* for the corresponding classical quantities. To study this analogy more closely let us see which quantity plays the role of the classical amplitude.

When the oscillator is in state n, the probability that coordinate ξ has a value lying between ξ and $\xi + d\xi$ is expressed by

$$|\psi_n(\xi)|^2 \, d\xi = \frac{1}{\pi^{1/2} 2^n n!} e^{-\xi^2} H_n^2(\xi) \, d\xi \qquad (8.13)$$

The asymptotic form of $\psi_n(\xi)$, (5.22), shows that for large values of n the function $\psi_n(\xi)$ is approximately a sinusoid when ξ changes in the interval between $-(2n)^{1/2}$ and $(2n)^{1/2}$, with the polynomial $H_n(\xi)$ vanishing exactly p times. Outside this interval $\psi_n(\xi)$ begins to decrease rapidly because of the predominance of the exponential factor. It follows then that the probability density is noticeably nonzero only in the interval $-(2n)^{1/2} < \xi < (2n)^{1/2}$, which implies that

$$\xi_0 = (2n)^{1/2} \qquad (8.14)$$

can be considered the "amplitude" of the oscillator. In conventional units the amplitude is

$$x_0 = \left(\frac{2n\hbar}{m\omega}\right)^{1/2} \qquad (8.15)$$

On the other hand, the energy of the oscillator is

$$E = E_n = \left(n + \frac{1}{2}\right) \hbar\omega \approx n\hbar\omega \qquad (8.16)$$

Solving (8.15) for n and substituting the result into (8.16), we get

$$E \approx \frac{1}{2} m\omega^2 x_0^2 \qquad (8.17)$$

Hence the relation between amplitude and energy is the same as in the classical theory. If we compare the matrix elements of $x(t)$ with the amplitude (8.15), we find that

$$\langle n| x(t) | n-1\rangle + \langle n-1 | x(t) |n\rangle = \left(\frac{n\hbar}{2m\omega}\right)^{1/2} (e^{i\omega t} + e^{-i\omega t})$$

$$= \left(\frac{2n\hbar}{m\omega}\right)^{1/2} \cos \omega t$$

$$= x_0 \cos \omega t \qquad (8.18)$$

Thus the matrix elements of $x(t)$ that are closest to the nth diagonal element give the terms in the Fourier series that repre-

sents the classical quantity $x(t)$ in the nth state (that is, in a state with $E = E_n$); here the values of n are considered large. This formal analogy between the terms in a Fourier series and the elements of a matrix representing a quantum operators served Heisenberg as a starting point 'for the creation of quantum mechanics, which in its earliest version was called *matrix mechanics*, as we know.

9. An elementary criterion for the applicability of the formulas of classical mechanics

When in Section 15, Chapter III, we considered the semiclassical approximation to the solution of the Schrödinger equation, we found the approximate expression of the wave function, ψ, in terms of the action integral S. Let us now apply the results of that section to the stationary state.

The Schrödinger equation in this case is

$$\nabla^2 \psi + \frac{2m}{\hbar^2} (E - U) \psi = 0 \tag{9.1}$$

and the Hamilton-Jacobi equation of classical mechanics reads

$$\frac{1}{2m} (\text{grad } S)^2 + U = E \tag{9.2}$$

If S is the general solution to Eq. (9.2), a solution that contains three arbitrary constants c_1, c_2, and c_3 (including the constant energy E but not the additive term), then depending on the boundary conditions we can put

$$\psi = \left(\det \left[\frac{\partial^2 S}{\partial x_i \, \partial x_k} \right] \right)^{1/2} e^{iS/\hbar} \tag{9.3}$$

or

$$\psi = \left| \left(\det \left[\frac{\partial^2 S}{\partial x_i \, \partial x_k} \right] \right)^{1/2} \right| \cos \left(\frac{S}{\hbar} + \alpha \right) \tag{9.4}$$

where the expression in parentheses is the determinant of the second derivatives of S, and α is a constant phase. The transition from the equation of wave mechanics, (9.1), to the equation of classical mechanics, (9.2), is formally analogous to the transformation from wave optics to geometrical optics. We can use the language of wave optics (or wave mechanics) to formulate the conditions under which the approximate formulas (9.3) and (9.4) can be applied:

The relative change in the index of refraction (or wavelength) over a distance of one wavelength must be considerably smaller than unity.

If instead of the wavelength λ we use $\lambda/(2\pi)$ as the characteristic length, this condition can be written as

$$\frac{\lambda}{2\pi} \times \frac{|\operatorname{grad}\lambda|}{\lambda} \equiv \left|\operatorname{grad}\frac{\lambda}{2\pi}\right| \ll 1 \qquad (9.5)$$

In quantum mechanics λ is the *de Broglie wavelength:*

$$\lambda = \frac{2\pi\hbar}{[2m\,(E-U)]^{1/2}} \qquad (9.6)$$

But since the domain we are considering lies between classical mechanics and quantum mechanics, the criterion for applying (9.3) or (9.4) can be stated in the language of classical mechanics as well. Indeed, if we substitute the expression for λ, (9.6), into condition (9.5), we get

$$\frac{m\hbar}{[2m\,(E-U)]^{3/2}} |\operatorname{grad}U| \ll 1 \qquad (9.7)$$

We introduce the absolute value of the particle's velocity, v, and the absolute value of its acceleration, w, and write

$$[2m\,(E-U)]^{1/2} = mv \qquad (9.8)$$

$$|\operatorname{grad}U| = mw \qquad (9.9)$$

Hence (9.7) gives

$$\frac{\hbar w}{mv^3} \ll 1 \qquad (9.10)$$

or

$$\frac{mv^3}{\hbar w} \gg 1 \qquad (9.11)$$

This is the very criterion that we were looking for. Apart from $\hbar$ it includes only quantities of classical mechanics; namely, kinematic quantities and the particle's mass.

Note that, according to the well-known formula of kinematics, we have

$$w^2 = \left(\frac{dv}{dt}\right)^2 + \left(\frac{v^2}{\rho}\right)^2 \qquad (9.12)$$

where v is the absolute value of the velocity, and ρ is the path's radius of curvature. This implies that

$$w \geqslant \frac{v^2}{\rho} \qquad (9.13)$$

which, after being substituted into (9.10), yields

$$\frac{\hbar}{mv\rho} \ll 1 \qquad (9.14)$$

or

$$\frac{\lambda}{2\pi\rho} \ll 1 \qquad (9.15)$$

where λ is again the de Broglie wavelength. Thus the de Broglie wavelength must be considerably shorter than the path's radius of curvature.

The criterion expressed by formulas (9.10) and (9.11) can be applied in two different ways. First, if we consider velocity and acceleration as functions of position coordinates [Eqs. (9.8) and (9.9)], then (9.3) or (9.4) will give a good approximation to Schrödinger's wave function. Second, instead of velocity and acceleration we can introduce their mean values into (9.13). The left-hand side will then become a certain constant whose order of magnitude will, when compared to unity, characterize the applicability of classical equations.

In the initial stages of development of quantum mechanics Bohr formulated his famous *correspondence principle*. According to this principle, for large quantum numbers the formulas of quantum mechanics must transform into classical formulas. We can then expect that the above-mentioned parameter [the left-hand side of (9.11)] will be connected with the quantum number characteristic of the given problem.

We will now show, using a simple example, that this is indeed the case. Let us consider the one-dimensional harmonic oscillator. Here the velocity is parallel to the acceleration. We will use the root-mean-square values of velocity and acceleration as the parameters. We have

$$x = a \cos \omega t \tag{9.16}$$

and hence

$$v^2 = \overline{\dot{x}^2} = \frac{1}{2} a^2 \omega^2, \qquad w^2 = \overline{\ddot{x}^2} = \frac{1}{2} a^2 \omega^4 \tag{9.17}$$

This implies that

$$\frac{mv^3}{\hbar w} = \frac{ma^2 \omega}{2\hbar} \tag{9.18}$$

The oscillator's energy is expressed in terms of the amplitude as

$$E = \frac{1}{2} ma^2 \omega^2 \tag{9.19}$$

and hence (9.18) reads

$$\frac{ma^2 \omega}{2\hbar} = \frac{E}{\hbar \omega} \tag{9.20}$$

Thus for the one-dimensional oscillator criterion (9.11) takes on the following form:

$$\frac{mv^3}{\hbar w} = \frac{E}{\hbar \omega} \gg 1 \tag{9.21}$$

But according to (8.16) the energy of the oscillator is

$$E = E_n = \left(n + \frac{1}{2} \right) \hbar \omega \tag{9.22}$$

where n is the quantum number of the oscillator in a given state. This implies that (9.21) demands that n be large compared to unity.

In summary we may say that the applicability of equations of the classical theory does not necessarily mean the applicability of classical concepts. The difference between the "probabilistic" way of describing phenomena, using a wave function, and the "absolute" way, using classical quantities, (this difference was mentioned in Part I) remains valid as well when classical quantities provide a good approximation to the wave function.

Chapter II

PERTURBATION THEORY

1. Statement of the problem

Only in a limited number of the simplest cases is it possible to solve the Schrödinger equation and find the eigenfunctions of the Hamiltonian, as well as of other operators. Solution of more complicated problems, say, the many-body problem, requires approximate, that is, essentially new, methods (for instance, the variational method). The many-electron problem will be studied in Part IV. Here we will consider the case when the process of solving a problem can be divided into two steps. The first step is to simplify the problem and then solve it exactly. The second is to calculate those *corrections* to the simplified problem that enable us to estimate the influence of the small terms omitted in the first step. There is a general method for calculating the corrections called *perturbation theory*, which we will develop in this chapter.

Let us assume that there exists an operator (for definiteness we consider it to be the Hamiltonian) that can be represented as a sum of two terms:

$$H = H^0 + \varepsilon U \qquad (1.1)$$

The first term, H^0, is the Hamiltonian of the *unperturbed problem*, or the *zeroth-order Hamiltonian*, and the second, εU, is the correction, or the *perturbation Hamiltonian*, which we can regard as "small" (for the sake of convenience we write the correction as the product of the *smallness parameter*, ε, by the operator, U). We will consider the perturbation Hamiltonian to be such that when ε is reduced to zero, the eigenvalues and eigenfunctions of H continuously transform to the eigenvalues and eigenfunctions of H^0.

This condition in some cases is not satisfied, and the perturbation changes the type of solution, introducing a continuous spectrum, for instance. The formal solution of the problem in such cases, however, has a physical meaning. The solution provides a wave function that describes the state of an atom that is not exactly stationary but can be considered *quasi-stationary*. But what is such a state? If we assume the obtained wave function to represent some initial state of an atom, then for a long period the state will but slightly differ from the initial one. The theory of quasi-stationary states will be discussed in Section 8, Chapter III.

We must bear in mind that the perturbation theory series can diverge, which does not, however, deprive it of physical meaning. This is so if the first terms in the series decrease relatively fast. In this case we use only a finite number of terms.

To return to our problem let us assume that the eigenfunctions $\psi_n^0(x)$ and eigenvalues E_n^0 of H^0 are known exactly, that is, we have found the solutions to

$$H^0\psi_n^0(x) = E_n^0\psi_n^0(x) \tag{1.2}$$

We need to find the approximate expressions for the eigenfunctions and eigenvalues of H, that is, solve the equation

$$(H^0 + \varepsilon U)\psi_n(x) = E_n\psi_n(x) \tag{1.3}$$

To solve the posed problem we must first solve the nonhomogeneous equation

$$H^0\psi - E'\psi = f \tag{1.4}$$

for the case when the parameter E' is one of the eigenvalues of H^0. Depending on whether the eigenvalues of H^0 are degenerate or nondegenerate, the solutions of both the preliminary problem and the general problem differ. To clarify the idea of the method we will first consider nondegenerate eigenvalues. After that we will generalize the results and apply them to the degenerate case.

2. Solution of the nonhomogeneous equation

Let us consider the nonhomogeneous equation

$$H^0\psi - E'\psi = f \tag{2.1}$$

where f is a known function, and ψ is the sought function. Assume E' to be one of the eigenvalues E_n^0 of H^0. Equation (2.1) then has the form

$$H^0\psi - E_n^0\psi = f \tag{2.2}$$

First we consider the eigenvalue E_n^0 to be *nondegenerate*, which means that the corresponding homogeneous equation

$$H^0\psi - E_n^0\psi = 0 \tag{2.3}$$

has only one solution, $\psi = \psi_n^0$.

We expand the unknown function, f, in a ψ_n^0-series:

$$f = \sum_m a_m\psi_m^0 + \int a(E)\,\psi_E^0\,dE \tag{2.4}$$

and look for the solution of (2.2) in the form of a similar series:

$$\psi = \sum_m c_m \psi_m^0 + \int c\,(E)\,\psi_E^0\,dE \qquad (2.5)$$

which, after substituting (2.4) and (2.5) into (2.2), yields

$$\sum_m c_m \left(E_m^0 - E_n^0\right) \psi_m^0 + \int c\,(E)\,(E - E_n^0)\,\psi_E^0\,dE$$

$$= \sum_m a_m \psi_m^0 + \int a\,(E)\,\psi_E^0\,dE \qquad (2.6)$$

In the left-hand side of this equation the coefficient of ψ_n^0 is zero. For the equation to have a (nontrivial) solution the coefficient of the corresponding term in the right-hand side must also vanish, that is

$$a_n = 0 \qquad (2.7)$$

which can be rewritten as

$$\int \bar{\psi}_n^0 f\,d\tau = 0 \qquad (2.8)$$

Thus we conclude that

> *The nonhomogeneous equation (2.2) has a solution if its right-hand side is orthogonal to the solution of the corresponding homogeneous equation.*

If this is so, the other coefficients c_m and $c(E)$ can be found by equating the corresponding terms on both sides of (2.6). This yields

$$c_m = \frac{a_m}{E_m^0 - E_n^0}, \qquad c\,(E) = \frac{a\,(E)}{E - E_n^0} \qquad (2.9)$$

Expansion (2.5) then reads

$$\psi = \sum_m{}' \frac{a_m}{E_m^0 - E_n^0}\,\psi_m + \int \frac{a\,(E)}{E - E_n^0}\,\psi_E^0\,dE \qquad (2.10)$$

where the prime on the summation sign means that the term with $m = n$ must be omitted.

To this expression obviously we can add a solution of the homogeneous equation, $c\psi_n^0$, where c is an arbitrary constant.

If in Eq. (2.1) E' had not been equal to any E_n^0, there would be no need to impose restrictions of type (2.8) on f, and the solution would have been

$$\psi = \sum_m \frac{a_m}{E_m^0 - E'}\,\psi_m^0 + \int \frac{a\,(E)}{E - E'}\,\psi_E^0\,dE \qquad (2.11)$$

where m runs through all values.

Now, let us turn to the *degenerate* case. As usual, we will consider E_0^0, E_1^0, E_2^0, ... to be different eigenvalues, so that their degeneracy will be expressed in the fact that each of them (say E_n^0) can have several (say s) eigenfunctions, which we denote

$$\psi_{n1}^0, \ \psi_{n2}^0, \ \ldots, \ \psi_{ns}^0 \qquad (2.12)$$

with s depending on n in the general case. We note that the eigenvalues corresponding to a continuous spectrum can also be degenerate. For the sake of simplicity, however, we will write our equations as though these eigenvalues were nondegenerate.

Let us assume that the homogeneous equation

$$H^0\psi_n^0 - E_n^0\psi_n^0 = 0 \qquad (2.3)$$

has s solutions (2.12) and we must find the solutions for the nonhomogeneous equation

$$H^0\psi - E_n^0\psi = f \qquad (2.2)$$

The expansions for the given function f and the unknown function ψ are then

$$f = \sum_m \sum_{r=1}^s a_{mr}\psi_{mr}^0 + \int a(E)\,\psi_E^0\,dE \qquad (2.13)$$

$$\psi = \sum_m \sum_{r=1}^s c_{mr}\psi_{mr}^0 + \int c(E)\,\psi_E^0\,dE \qquad (2.14)$$

Substituting these into (2.2), we get

$$\sum_m (E_m^0 - E_n^0) \sum_{r=1}^s c_{mr}\psi_{mr}^0 + \int (E - E_n^0)\,c(E)\,\psi_E^0\,dE$$

$$= \sum_m \sum_{r=1}^s a_{mr}\psi_{mr}^0 + \int a(E)\,\psi_E^0\,dE \qquad (2.15)$$

From this we conclude that

$$a_{n1} = a_{n2} = \ldots = a_{ns} = 0 \qquad (2.16)$$

that is, f must satisfy s conditions

$$\int \bar{\psi}_{nr}^0 f\,d\tau = 0, \qquad r = 1, 2, \ldots, s \qquad (2.17)$$

Thus we see that in the degenerate case

The nonhomogeneous equation (2.2) *has a solution if its right-hand side is orthogonal to each solution of the corresponding homogeneous equation.*

After we have determined all the coefficients c_m and $c(E)$, we get the following expression for ψ:

$$\psi = \sum_m{}' \frac{1}{E_m^0 - E_n^0} \sum_{r=1}^{s} a_{mr}\psi_{mr}^0 + \int \frac{a(E)}{E - E_n^0} \psi_E^0 \, dE \qquad (2.18)$$

We see that for the degenerate and the nondegenerate case the equations are quite analogous and the criterion for the existence of a solution is formulated in almost the same way in both cases.

3. Nondegenerate eigenvalues

Let us now return to our main problem, namely, the solution of

$$(H^0 + \varepsilon U)\,\psi_n = E_n\psi_n \qquad (3.1)$$

and study the case when all the eigenvalues of H^0 are nondegenerate.

We look for the eigenvalue E_n and the eigenfunction ψ_n in the form of expansions in powers of the smallness parameter ε:

$$E_n = E_n^0 + \varepsilon E_n' + \varepsilon^2 E_n'' + \ldots \qquad (3.2)$$

$$\psi_n = \psi_n^0 + \varepsilon\psi_n' + \varepsilon^2\psi_n'' + \ldots \qquad (3.3)$$

If we substitute these into Eq. (3.1) and identify powers of ε, we get a series of equations

$$H^0\psi_n^0 - E_n^0\psi_n^0 = 0 \qquad (3.4a)$$

$$H^0\psi_n' - E_n^0\psi_n' = -U\psi_n^0 + E_n'\psi_n^0 \qquad (3.4b)$$

$$H^0\psi_n'' - E_n^0\psi_n'' = -U\psi_n' + E_n'\psi_n' + E_n''\psi_n^0 \qquad (3.4c)$$

. . . .

The first one is satisfied automatically since, according to the initial assumption, ψ_n^0 is an eigenfunction of H^0 corresponding to E_n^0. The second equation, (3.4b), is a nonhomogeneous equation for determining ψ_n'. As we know, for this equation to have a solution its right-hand side must be orthogonal to the solution of the corresponding homogeneous equation, ψ_n^0. Using the fact that ψ_n^0 is normalized, we can write this condition as

$$E_n' = (n \mid U \mid n) \qquad (3.5)$$

where

$$(n \mid U \mid n) = \int \bar{\psi}_n^0 U\psi_n^0 \, d\tau \qquad (3.6)$$

is the *diagonal element* of the matrix for the perturbation U. Thus the orthogonality condition makes it possible to determine the

unknown constant E'_n. The next step in solving (3.4b) is to expand its right-hand side in terms of the functions ψ^0_m and ψ^0_E:

$$f = -U\psi^0_n + E'_n\psi^0_n$$

$$= -\sum_m{}' (m\,|\,U\,|\,n)\,\psi^0_m - \int (E\,|\,U\,|\,n)\,\psi^0_E\,dE \qquad (3.7)$$

where

$$(m\,|\,U\,|\,n) = \int \bar\psi^0_m U\psi^0_n\,d\tau \qquad (3.8)$$

$$(E\,|\,U\,|\,n) = \int \bar\psi^0_E U\psi^0_n\,d\tau \qquad (3.8^*)$$

The solution ψ'_n of Eq. (3.4b) can be obtained by using the formulas of the previous section, namely

$$\psi'_n = \sum_m{}' \frac{(m\,|\,U\,|\,n)}{E^0_n - E^0_m}\,\psi^0_m + \int \frac{(E\,|\,U\,|\,n)}{E^0_n - E}\,\psi^0_E\,dE \qquad (3.9)$$

(We did not add the term $c\psi^0_n$ because ψ'_n is obviously orthogonal to ψ^0_n. Hence

$$\psi_n = \psi^0_n + e\psi'_n \qquad (3.10)$$

which is the approximate solution to the perturbation problem, will be normalized to within terms of the order of ε^2_n.)

Let us now turn to the third equation, (3.4c). First, we must determine the constant E''_n from the condition that the equation has a solution. This condition yields

$$E''_n = \int \bar\psi^0_n \left(U - E'_n\right) \psi'_n\,d\tau \qquad (3.11)$$

If we then substitute expansion (3.9) for ψ'_n, from the completeness property of eigenfunctions we get

$$E''_n = \sum_m{}' \frac{|\,(m\,|\,U\,|\,n)\,|^2}{E^0_n - E^0_m} + \int \frac{|\,(E\,|\,U\,|\,n)\,|^2}{E^0_n - E}\,dE \qquad (3.12)$$

After this we could find the correction ψ''_n and the third, fourth, etc. approximations. The calculations proceed in a similar manner, notably, after we find

$$\psi^0_n,\ \psi'_n,\ \ldots,\ \psi^{(k-1)}_n \quad \text{and} \quad E^0_n,\ E'_n,\ E''_n,\ \ldots,\ E^{(k-1)}_n$$

the condition for the existence of a solution to the kth equation yields $E^{(k)}_n$ and then $\psi^{(k)}_n$. The formulas become more and more involved, but it is usually sufficient to take the first-order approximation for the eigenfunction and the second-order approximation

for the eigenvalue. When E_n', the first-order correction for the eigenvalue E_n, is nonzero, we can limit ourselves to this if there is no need for a higher accuracy.

4. Degenerate eigenvalues. Expansion in powers of the smallness parameter

Let us look for the equation

$$(H^0 + \varepsilon U)\, \psi_n = E_n \psi_n \tag{4.1}$$

for the case when there is *degeneracy*, that is, eigenvalues of H^0 are degenerate. We choose the unperturbed equation

$$H^0 \psi_{nr}^0 - E_n^0 \psi_{nr}^0 = 0 \tag{4.2}$$

to have s solutions

$$\psi_{n1}^0, \ \psi_{n2}^0, \ \ldots, \ \psi_{ns}^0 \tag{4.3}$$

We know (Section 7, Chapter II, Part I) that one can select the s solutions to a certain extent arbitrarily, which implies that by performing a unitary transformation we can replace the functions (4.3) by a linear combination of them. For what follows it is convenient to construe the ψ_{nr}^0 as solutions chosen in a specific way depending on U. The original s solutions (whatever they may be), whose linear combinations are ψ_{nr}^0, we will denote

$$\varphi_1, \ \varphi_2, \ \ldots, \ \varphi_s \tag{4.4}$$

where we omit the subscript n though keeping it in mind.

We know from algebra that in an infinitesimal change of the coefficients of an algebraic equation a multiple root can separate into several simple roots. Here too by analogy we can expect a degenerate eigenvalue E_n^0 to separate into s nondegenerate eigenvalues E_{nr} ($r = 1, 2, \ldots, s$) owing to the perturbation. For this reason we will seek the eigenvalues in the form

$$E_{nr} = E_n^0 + \varepsilon E_{nr}' + \varepsilon^2 E_{nr}'' + \ldots \tag{4.5}$$

where the corrections depend on the number r of the corresponding eigenfunction. This last can be written as

$$\psi_{nr} = \psi_{nr}^0 + \varepsilon \psi_{nr}' + \varepsilon^2 \psi_{nr}'' + \ldots \tag{4.6}$$

By substituting (4.5) and (4.6) into (4.1) (where we must change E_n to E_{nr} and ψ_n to ψ_{nr}) and identifying powers of ε we get a series of equations

$$H^0\psi_{nr}^0 - E_n^0\psi_{nr}^0 = 0 \qquad (4.7a)$$

$$H^0\psi_{nr}' - E_n^0\psi_{nr}' = -U\psi_{nr}^0 + E_{nr}'\psi_{nr}^0 \qquad (4.7b)$$

$$H^0\psi_{nr}'' - E_n^0\psi_{nr}'' = -U\psi_{nr}' + E_{nr}'\psi_{nr}' + E_{nr}''\psi_{nr}^0 \qquad (4.7c)$$

$$\cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot$$

These differ from the analogous equations (3.4a)—(3.4c) only in the addition of a second subscript to the eigenvalues and eigenfunctions. The first equation, (4.7a), is satisfied automatically. For the second equation, (4.7b), to have a solution its right-hand side must be orthogonal to each solution of the homogeneous equation:

$$\int \bar{\psi}_{np}^0 (U - E_{nr}') \psi_{nr}^0 \, d\tau = 0, \qquad p = 1, 2, \ldots, s \qquad (4.8)$$

or, which is the same,

$$\int \bar{\varphi}_p^0 (U - E_{nr}') \psi_{nr}^0 \, d\tau = 0, \qquad p = 1, 2, \ldots, s \qquad (4.9)$$

Equations (4.8) and (4.9) are not generally satisfied by arbitrary solutions of (4.2). So we must find combinations

$$\psi_{nr}^0 = b_{1r}\varphi_1 + b_{2r}\varphi_2 + \ldots + b_{sr}\varphi_s \qquad (4.10)$$

of the known solutions (4.4) that will satisfy (4.8) and (4.9).

5. The eigenfunctions in the zeroth-order approximation

Let us find the coefficients in the unitary transformation (4.10). If we substitute (4.10) into (4.9) and recall that the functions φ_q are orthogonal and normalized, we get

$$\sum_{q=1}^{s} U_{pq}b_{qr} = E_{nr}'b_{pr} \qquad (5.1)$$

where

$$U_{pq} = \int \bar{\varphi}_p U\varphi_q \, d\tau \qquad (5.2)$$

Omitting the second subscript in b_{qr} and denoting the unknown quantity E_{nr}' by λ, we can write (5.1) as

$$\sum_{q=1}^{s} U_{pq}b_q = \lambda b_q \qquad (5.3)$$

These equations can be interpreted as the equations for the eigenfunctions $b(q) = b_q$ of the operator represented by the finite matrix U_{pq}. This operator is hermitian since its matrix is hermitian, which can be seen from (5.2). This implies that its eigenvalues λ are real. To find these and solve Eq. (5.3) we can use a purely algebraic method. Equation (5.3) is a system of homogeneous linear equations in b_q. Such a system will have a unique solution if the system determinant

$$D(\lambda) = \begin{vmatrix} U_{11} - \lambda & U_{12} & \ldots & U_{1s} \\ U_{21} & U_{22} - \lambda & \ldots & U_{2s} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ U_{s1} & U_{s2} & \ldots & U_{ss} - \lambda \end{vmatrix} \tag{5.4}$$

is zero. The equation $D(\lambda) = 0$ has s real roots, and each root $\lambda = \lambda_r$ has a corresponding solution $b_q = b_{qr}$ of (5.3). These solutions can be normalized so that

$$|b_1|^2 + |b_2|^2 + \ldots + |b_s|^2 = 1 \tag{5.5}$$

It follows from the general properties of linear operators that solutions

$$b_q = b_{qr} \quad \text{and} \quad b_q = b_{qp}$$

corresponding to two distinct roots

$$\lambda = \lambda_r \quad \text{and} \quad \lambda = \lambda_p$$

are orthogonal to each other, that is

$$\bar{b}_{1r}b_{1p} + \bar{b}_{2r}b_{2p} + \ldots + \bar{b}_{sp}b_{sr} = 0, \qquad \lambda_r \neq \lambda_p \tag{5.6}$$

The equation $D(\lambda) = 0$ can also have multiple roots. For instance, if $\lambda = \lambda_r$ is a double root, there are two independent solutions of Eq. (5.3), $b_q = b'_q$ and $b_q = b''_q$, corresponding to this root. But no matter whether $D(\lambda) = 0$ has simple or multiple roots we can always make all the s solutions of (5.3) orthogonal and normalized, so that

$$\bar{b}_{1p}b_{1q} + \bar{b}_{2p}b_{2q} + \ldots + \bar{b}_{sp}b_{sq} = \delta_{pq} \tag{5.7}$$

which implies that matrix b with elements b_{pq} is unitary. Indeed, if we assume, as usual, that

$$\bar{b}_{qp} = b^{+}_{pq} \tag{5.8}$$

we can write (5.7) as

$$\sum_{r=1}^{s} b^{+}_{pr}b_{rq} = \delta_{pq} \tag{5.9}$$

It then follows, and this can easily be proved, that

$$\sum_{r=1}^{s} b_{pr}b_{rq}^{+} = \delta_{pq} \tag{5.10}$$

In matrix notation we may write (5.9) and (5.10), which express the unitarity of matrix b, in the form

$$b^{+}b = 1, \qquad bb^{+} = 1 \tag{5.11}$$

We have thus found the unitary transformation (4.10) whose coefficients satisfy Eq. (5.1), where the E'_{nr} are the roots λ_r of equation $D(\lambda) = 0$:

$$E'_{nr} = \lambda_r, \qquad D(\lambda_r) = 0 \tag{5.12}$$

The unitarity of transformation (4.10) shows that if the φ_p are orthogonal and normalized, ψ^0_{nr} will be too.

We must note that the ψ^0_{nr} are determined uniquely only if all roots of $D(\lambda) = 0$ are distinct. But if $\lambda = \lambda_1$ is, say, a double root, then instead of the solutions $b_{q1} = b'_q$ and $b_{q2} = b''_q$ of Eq. (5.3) we could have taken

$$b_{q1}^{*} = v_{11}b_{q1} + v_{21}b_{q2}$$
$$b_{q2}^{*} = v_{12}b_{q1} + v_{22}b_{q2} \tag{5.13}$$

where the matrix made up of the coefficients v_{ik} is unitary. The new solution (5.13) corresponds to new wave functions

$$\psi_{n1}^{*0} = v_{11}\psi_{n1}^{0} + v_{21}\psi_{n2}^{0}$$
$$\psi_{n2}^{*0} = v_{12}\psi_{n1}^{0} + v_{22}\psi_{n2}^{0} \tag{5.14}$$

Hence for each multiple root there is an arbitrary unitary transformation of the functions corresponding to this root. Such unitary transformations, which remain arbitrary in the first approximation, are usually determined from the second and higher approximations.

We can express the roots $E'_{nr} = \lambda_r$ in terms of the ψ^0_{nr}. If we denote

$$(nr\,|\,U\,|\,n'r') = \int \bar{\psi}^0_{nr} U \psi^0_{n'r'}\, d\tau \tag{5.15}$$

then on the basis of (4.8) we get

$$(nr\,|\,U\,|\,n'r') = E'_{nr}\delta_{rr'} \tag{5.16}$$

which at $r = r'$ yields the sought for expression for E'_{nr}.

6. The first and higher approximations

We now turn to Eq. (4.7b). With our choice of ψ_{nr}^0 its right-hand side satisfies the condition for the existence of a solution. The expansion of the right-hand side then reads

$$- U\psi_{nr}^0 + E_{nr}'\psi_{nr}^0 = - \sum_{mp}{}' (mp \,|\, U \,|\, nr)\,\psi_{mp}^0 - \int (E \,|\, U \,|\, nr)\,\psi_E^0 \, dE \quad (6.1)$$

where we have used (5.15) and put

$$(E \,|\, U \,|\, nr) = \int \bar{\psi}_E^0 U \psi_{nr}^0 \, d\tau \quad (6.2)$$

The prime on the summation sign in (6.1) means that the terms with $m = n$ must be omitted from the sum. After solving Eq. (4.7b) by the method of Section 2 we get

$$\psi_{nr}' = \sum_m{}' \frac{1}{E_n^0 - E_m^0} \sum_{p=1}^s (mp \,|\, U \,|\, nr)\, \psi_{mp}^0$$

$$+ \int \frac{(E \,|\, U \,|\, nr)}{E_n^0 - E}\, \psi_E^0 \, dE + \sum_{p=1}^s c_{pr}\psi_{np}^0 \quad (6.3)$$

Here the last sum is the solution of the homogeneous equation. The constants in it, c_{pr}, are unknown and must be found from the second approximation. In dealing with the second approximation, Eq. (4.7c), we must first make sure that its right-hand side is orthogonal to all solutions of the homogeneous equation. This condition reads

$$E_{nr}''\delta_{qr} = \int \bar{\psi}_{nq}^0 \,(U - E_{nr}')\, \psi_{nr}' \, d\tau \quad (6.4)$$

If now we substitute the expression (6.3) for ψ_{nr}' and introduce the notation

$$U_{qr}'' = \sum_m{}' \frac{1}{E_n^0 - E_m^0} \sum_{p=1}^s (nq \,|\, U \,|\, mp)\,(mp \,|\, U \,|\, nr)$$

$$+ \int \frac{(nq \,|\, U \,|\, E)\,(E \,|\, U \,|\, nr)}{E_n^0 - E}\, dE \quad (6.5)$$

we can write Eq. (6.4) as

$$E_{nr}''\delta_{qr} = U_{qr}'' + (E_{nq}' - E_{nr}')\, c_{qr} \quad (6.6)$$

where we have used (5.16). At $q \neq r$ this reduces to

$$U_{qr}'' + (E_{nq}' - E_{nr}')\, c_{qr} = 0, \qquad q \neq r \quad (6.7)$$

If all the E'_{nq} $(q = 1, 2, \ldots, s)$ are different, that is, all roots of $D(\lambda) = 0$ (see Section 5) are simple roots, then (6.7) gives us all the c_{qr} with different subscripts, namely

$$c_{qr} = - \frac{U''_{qr}}{E'_{nq} - E'_{nr}} \tag{6.8}$$

But if some roots are multiple, for instance, $E'_{n1} = E'_{n2}$, the corresponding sums $U''_{qr} = U''_{12}$ must vanish. This condition will be satisfied if we properly choose the unitary transformation (5.14), which up till now has remained arbitrary. Indeed, if we replace ψ^0_{n1} and ψ^0_{n2} by their linear combinations ψ^{*0}_{n1} and ψ^{*0}_{n2}, then U''_{qr} will change to

$$U^{*''}_{qr} = \sum_{i,\,k=1}^{2} v^+_{qi} U''_{ik} v_{kr}, \qquad q,\, r = 1,\, 2 \tag{6.9}$$

But according to (6.6) this must be equal to $E''_{nr}\delta_{qr}$:

$$\sum_{i,\,k=1}^{2} v^+_{qi} U''_{ik} v_{kr} = E''_{nr}\delta_{qr} \tag{6.10}$$

If we now premultiply this result by v_{pq} and sum with respect to q, we get

$$\sum_{k=1}^{2} U''_{pk} v_{kr} = E''_{nr} v_{pr} \tag{6.11}$$

We have come to equations similar to (5.1) and by using the method of Section 5 can find matrix $[v_{ik}]$. It could happen, however, that matrix $[v_{ik}]$ would be determined nonuniquely for the same reason as matrix $[b_{ik}]$ of Section 5. We would then have to bring in the higher approximations.

Let us assume that we have found $[v_{ik}]$ and that the functions ψ^0_{qr} are "modified", that is, if necessary, changed to their linear combinations. Then $E'_{nq} = E'_{nr}$ will satisfy Eq. (6.7) automatically and the corresponding c_{qr} will remain undefined. If $E'_{nq} \neq E'_{nr}$, we can find c_{qr} by using (6.8). If $q = r$, then according to Eq. (6.6) c_{rr} also remains undefined and can be nullified. Then the function $\psi_{nr} = \psi^0_{nr} + \varepsilon \psi'_{nr}$ will be normalized to within terms of the order of ε^2. Finally, the second-order energy correction, E''_{nr}, will be

$$E''_{nr} = U''_{rr} \tag{6.12}$$

where U''_{rr} is understood to be the modified U''_{rr}, which before we had denoted $U^{*''}_{rr}$.

We have obtained the first approximation to the eigenfunction ψ_{nr} and the second approximation to the energy. In a similar manner we could find the higher approximations, but in view of the complexity of the formulas they are of no practical interest.

7. The case of adjacent eigenvalues

The formulas of Section 3 show that if two eigenvalues E_n^0 and $E_{n'}^0$ of the zeroth-order Hamiltonian are *adjacent*, the denominators $E_n^0 - E_{n'}^0$ in expressions (3.9) and (3.12) for ψ_n' and E_n'' become small. Because of this the approximations will be poor, and if the denominators $E_n^0 - E_{n'}^0$ are of the same order of magnitude as the numerators (multiplied by ε), the above expressions cannot be used at all. In other words, the desired wave functions and eigenvalues cannot in this case be expanded in powers of the smallness parameter. Instead we can use a method in which the calculations are "reshuffled" so that in the terms with small denominators the numerators are zero. In dealing with this method we will study only the zeroth-order approximation.

We write the equation in question as

$$H\psi = E\psi \tag{7.1}$$

where

$$H = H^0 + \varepsilon U \tag{7.2}$$

and assume that the zeroth-order Hamiltonian, H^0, has two adjacent eigenvalues E_1^0 and E_2^0 which correspond to the eigenfunctions ψ_1^0 and ψ_2^0. We ask for that solution to (7.1) for which E is close to E_1^0 (and E_2^0). As the starting approximation we will take not ψ_1^0 or ψ_2^0 but their linear combination

$$\psi^* = c_1\psi_1^0 + c_2\psi_2^0 \tag{7.3}$$

as was done in Section 3. Indeed, from (3.9) we conclude that the leading term in ψ_1' is proportional to ψ_2^0. For this reason it is advisable to include this term in the starting approximation. Furthermore, from (4.10) it follows that in the limiting case, when E_1^0 coincides with E_2^0, the linear combination of ψ_1^0 and ψ_2^0 will be the zeroth-order approximation for the wave function. Substituting (7.3) into (7.1), we get the approximate equality

$$c_1(H - E)\psi_1 + c_2(H - E)\psi_2 = 0 \tag{7.4}$$

If we premultiply (7.4) first by $\bar\psi_1$, then by $\bar\psi_2$, and in both cases integrate, we come to two equations:

$$(H_{11} - E)c_1 + H_{12}c_2 = 0$$
$$H_{21}c_1 + (H_{22} - E)c_2 = 0 \tag{7.5}$$

9*

where

$$H_{ik} = \int \bar{\psi}_i^0 H \psi_k^0 \, d\tau, \qquad i, \, k = 1, \, 2 \qquad (7.6)$$

Notice that the off-diagonal matrix elements H_{12} and H_{21} are small compared to the diagonal elements H_{11} and H_{22}, since for the zeroth-order Hamiltonian $H_{12}^0 = H_{21}^0 = 0$. Equations (7.5) serve to determine the coefficients c_1 and c_2 and the parameter E. We nullify the determinant

$$D(E) = \begin{vmatrix} H_{11} - E & H_{12} \\ H_{21} & H_{22} - E \end{vmatrix} \qquad (7.7)$$

and solve the quadratic equation

$$E^2 - (H_{11} + H_{22})E + H_{11}H_{22} - |H_{12}|^2 = 0 \qquad (7.8)$$

to get for E two values

$$E = \frac{1}{2}(H_{11} + H_{22}) \pm \frac{1}{2}[(H_{11} - H_{22})^2 + 4|H_{12}|^2]^{1/2} \qquad (7.9)$$

which we denote E_1^* and E_2^*. It follows from (7.2) that

$$H_{ik} = E_k^0 \delta_{ik} + \varepsilon U_{ik} \qquad (7.10)$$

where U_{ik} is determined by a formula similar to (7.6). Therefore H_{12} is of the order of ε, and if $E_1^0 - E_2^0$ were not small, we could have expanded (7.9) in a power series in ε. But since this difference is small by definition, the expansion will either converge poorly or even diverge. It is now clear why the method of Section 3 does not work when two eigenvalues are adjacent. In the first-order approximation the two values of E given by (7.9) yield the perturbed eigenvalues E_1 and E_2 of H, the eigenvalues that correspond to the zeroth-order eigenvalues E_1^0 and E_2^0. We can use Eqs. (7.5) to find the coefficients c_1 and c_2. This is most conveniently done by introducing two auxiliary real quantities, α and β:

$$\frac{2H_{12}}{H_{11} - H_{22}} = -\tan\alpha\, e^{i\beta} \qquad (7.11)$$

From Eqs. (7.5), (7.9), and (7.11) it is easy to see that the ratio c_1/c_2 takes two values:

$$\left(\frac{c_1}{c_2}\right)_1 = \frac{c_{11}}{c_{21}} = -\cot\frac{\alpha}{2}\, e^{i\beta}$$

$$\left(\frac{c_1}{c_2}\right)_2 = \frac{c_{12}}{c_{22}} = \tan\frac{\alpha}{2}\, e^{i\beta} \qquad (7.12)$$

If we now choose

$$c_{11} = \cos\frac{\alpha}{2}\, e^{i\beta/2}, \qquad c_{12} = \sin\frac{\alpha}{2}\, e^{i\beta/2}$$

$$c_{21} = -\sin\frac{\alpha}{2}\, e^{-i\beta/2}, \qquad c_{22} = \cos\frac{\alpha}{2}\, e^{-i\beta/2} \tag{7.13}$$

we get the normalized solutions to (7.5). The corresponding eigenfunctions will be

$$\psi_1^* = \cos\frac{\alpha}{2}\, e^{i\beta/2}\psi_1^0 - \sin\frac{\alpha}{2}\, e^{-i\beta/2}\psi_2^0$$

$$\psi_2^* = \sin\frac{\alpha}{2}\, e^{i\beta/2}\psi_1^0 + \cos\frac{\alpha}{2}\, e^{-i\beta/2}\psi_2^0 \tag{7.14}$$

It is these functions that serve as the starting approximation. If we use them to compose the matrix elements of H, we get

$$H_{ik}^* = \int \bar\psi_i^* H\psi_k^*\, d\tau = E_k^* \delta_{ik}, \qquad i, k = 1, 2 \tag{7.15}$$

so that $H_{12}^* = 0$. Therefore, if we go on to the next approximation and build expressions similar to (3.9) or (3.12) using ψ_1^*, ψ_2^* and the remaining functions ψ_3^0, ψ_4^0, ..., we will not get terms with the small denominator $E_1^0 - E_2^0$. This implies that these expressions will indeed represent small corrections.

8. The anharmonic oscillator

As an example of how perturbation theory works we will take the case of the *one-dimensional anharmonic oscillator*. We assume that the system of units is such that the eigenvalue equation for the Hamiltonian is

$$-\frac{1}{2}\frac{d^2\psi}{d\xi^2} + \left(\frac{1}{2}\xi^2 + \varepsilon\xi^3\right)\psi = E\psi \tag{8.1}$$

where $\varepsilon\xi^3 = \varepsilon U$ is the perturbation Hamiltonian. We ask for the eigenfunction in the first-order approximation and the eigenvalue in the second. The eigenfunctions and eigenvalues of the unperturbed equation

$$-\frac{1}{2}\frac{d^2\psi_n^0}{d\xi^2} + \frac{1}{2}\xi^2\psi_n^0 = E_n^0\psi_n^0 \tag{8.2}$$

are already known. They are

$$\psi_n^0 = \pi^{-1/4}\left(2^n n!\right)^{-1/2} e^{-\xi^2/2} H_n(\xi) \tag{8.3}$$

$$E_n^0 = n + \frac{1}{2} \tag{8.4}$$

The perturbation operator, $U = \xi^3$, has matrix elements

$$(n \mid U \mid n') = \int\limits_{-\infty}^{+\infty} \bar{\psi}_n^0 \xi^3 \psi_{n'}^0 \, d\xi \qquad (8.5)$$

and the simplest way to obtain these elements is to multiply the matrices ξ and ξ^2 of Section 6, Chapter I, whose elements are

$$(n \mid \xi \mid n') = \left(\frac{n}{2}\right)^{1/2} \delta_{n-1\,n'} + \left(\frac{n+1}{2}\right)^{1/2} \delta_{n+1\,n'} \qquad (8.6)$$

$$(n \mid \xi^2 \mid n') = \frac{1}{2} [n(n-1)]^{1/2} \delta_{n-2\,n'}$$

$$+ \left(n + \frac{1}{2}\right) \delta_{nn'} + \frac{1}{2} [(n+1)(n+2)]^{1/2} \delta_{n+2\,n'} \qquad (8.7)$$

We find that

$$(n \mid U \mid n') = (n \mid \xi^3 \mid n') = \sum_k (n \mid \xi \mid k)(k \mid \xi^2 \mid n')$$

$$= \left(\frac{n(n-1)(n-2)}{8}\right)^{1/2} \delta_{n-3\,n'} + \left(\frac{9n^3}{8}\right)^{1/2} \delta_{n-1\,n'}$$

$$+ \left(\frac{9(n+1)^3}{8}\right)^{1/2} \delta_{n+1\,n'}$$

$$+ \left(\frac{(n+1)(n+2)(n+3)}{8}\right)^{1/2} \delta_{n+3\,n'} \qquad (8.8)$$

The diagonal element of U is zero, which can also be directly observed from formula (8.5). Hence the first order correction to the eigenvalue vanishes. The approximate expression for the eigenfunction is

$$\psi_n = \psi_n^0 + \varepsilon\psi_n' \qquad (8.9)$$

where according to (3.9)

$$\psi_n' = \sum_m{}' \frac{(m \mid U \mid n)}{E_n^0 - E_m^0} \psi_m^0 \qquad (8.10)$$

In our case only four terms in this sum are nonzero, namely

$$\psi_n' = \frac{(n-3 \mid U \mid n)}{E_n^0 - E_{n-3}^0} \psi_{n-3}^0 + \frac{(n-1 \mid U \mid n)}{E_n^0 - E_{n-1}^0} \psi_{n-1}^0$$

$$+ \frac{(n+1 \mid U \mid n)}{E_n^0 - E_{n+1}^0} \psi_{n+1}^0 + \frac{(n+3 \mid U \mid n)}{E_n^0 - E_{n+3}^0} \psi_{n+3}^0 \qquad (8.10^*)$$

Substituting the matrix elements according to (8.8), we finally get

$$\psi'_n = \frac{1}{3}\left(\frac{n(n-1)(n-2)}{8}\right)^{1/2}\psi^0_{n-3} + 3\left(\frac{n^3}{8}\right)^{1/2}\psi^0_{n-1}$$

$$- 3\left(\frac{(n+1)^3}{8}\right)^{1/2}\psi^0_{n+1}$$

$$- \frac{1}{3}\left(\frac{(n+1)(n+2)(n+3)}{8}\right)^{1/2}\psi^0_{n+3} \quad (8.11)$$

We still have to calculate, using (3.12), the second-order correction to the energy. We find that

$$E''_n = \frac{1}{3}\frac{n(n-1)(n-2)}{8} + \frac{9}{8}n^3 - \frac{9}{8}(n+1)^3$$

$$- \frac{1}{3}\frac{(n+1)(n+2)(n+3)}{8}$$

$$= -\frac{15}{4}\left(n^2 + n + \frac{11}{30}\right) \quad (8.12)$$

Hence the energy in this approximation is

$$E_n = E^0_n + \varepsilon^2 E''_n = \left(n + \frac{1}{2}\right) - \frac{15}{4}\varepsilon^2\left(n^2 + n + \frac{11}{30}\right) \quad (8.13)$$

and we have completed the solution.

Now let us assume that besides $\varepsilon\xi^3$ the perturbation Hamiltonian has a term $\delta\xi^4$. The eigenvalue equation will then be

$$-\frac{1}{2}\frac{d^2\psi}{d\xi^2} + \left(\frac{1}{2}\xi^2 + \varepsilon\xi^3 + \delta\xi^4\right)\psi = E\psi \quad (8.14)$$

If δ is of the order of ε^2, the first-order approximation for the eigenfunction will remain unchanged and to the expression (8.13) for the eigenvalue a new term will be added, the diagonal matrix element of $\delta\xi^4$. Let us calculate the additional term. We have

$$(n\,|\,\xi^4\,|\,n) = (n\,|\,\xi\,|\,n-1)(n-1\,|\,\xi^3\,|\,n)$$

$$+ (n\,|\,\xi\,|\,n+1)(n+1\,|\,\xi^3\,|\,n)$$

$$= \frac{3}{4}n^2 + \frac{3}{4}(n+1)^2$$

so that

$$\delta\,(n\,|\,\xi^4\,|\,n) = \frac{3}{2}\delta\left(n^2 + n + \frac{1}{2}\right) \quad (8.15)$$

Adding (8.15) to (8.13) yields the eigenvalue E_n:

$$E_n = \left(n + \frac{1}{2}\right) + \left(\frac{3}{2}\delta - \frac{15}{4}\varepsilon^2\right)\left(n + \frac{1}{2}\right)^2 + \frac{3}{8}\delta - \frac{7}{16}\varepsilon^2 \quad (8.16)$$

In the particular case of $\delta = 5\varepsilon^2/2$ formula (8.16) differs from E_n^0 only by a constant term $\varepsilon^2/2$.

We note in conclusion that by adding terms of type $\varepsilon\xi^3$ or $\delta\xi^4$ to the potential energy of the harmonic oscillator we have changed the type of eigenfunction of the Hamiltonian. We have thus come to a case mentioned in Section 1, when the formal application of perturbation theory leads to divergent series and gives a state that is not exactly stationary but quasi-stationary. The formal character of the solution appears in the fact that for large values of n the correction to the energy in (8.16) ceases to be small.

Chapter III

RADIATION, THE THEORY OF DISPERSION, AND THE LAW OF DECAY

1. Classical formulas

The quantum laws governing phenomena in which the finiteness of the speed of propagation of actions (interaction) does not play any role have clearly been established. They can be formulated in terms of the concepts introduced in Part I and also by bringing in a new principle (the famous *Pauli exclusion principle*) needed for statement of the many-body problem. We will examine Pauli's theory of the electron and the many-body problem in Parts III and IV.

On the other hand, the theory of phenomena in which the finiteness of the speed of propagation of actions does play a role has yet to be completed. These phenomena include, first of all, those that are studied in electrodynamics and the theory of relativity.

Quantum generalization of each of these two theories requires considering systems of an indefinite number of particles (photons in the case of electrodynamics and electrons and positrons in relativistic quantum mechanics). Here the full theory of such systems will not be given, nor will the quantum theory of radiation (quantum electrodynamics). We will confine ourselves to the derivation of formulas using the classical theory as guide. In doing so we will allow ourselves a certain degree of inconsistency. Namely, we will describe the atomic system in quantum terms and the radiation in classical terms, not directly introducing the concept of light quanta (photons). Aside from this, the approximate (semiclassical) nature of the theory can serve to justify a certain inprecision of expression, such as "the electron is inside a given volume" (instead of "the electron can be detected in a given volume by a certain kind of experiment").

Before considering the theory of radiation let us recall the main formulas of the classical theory. The Maxwell equations in Lorentz form are

$$\text{curl } \mathscr{H} - \frac{1}{c} \frac{\partial \mathscr{E}}{\partial t} = \frac{4\pi}{c} \rho \mathbf{v}$$

$$\text{div } \mathscr{E} = 4\pi\rho \tag{1.1}$$

$$\text{curl } \mathscr{E} + \frac{1}{c} \frac{\partial \mathscr{H}}{\partial t} = 0$$

$$\text{div } \mathscr{H} = 0 \tag{1.2}$$

Here $\mathscr{E}$ and $\mathscr{H}$ are the *electric* and *magnetic field vectors*, ρ is the electric charge density, $\mathbf{v}$ is the velocity of the electrons, and c is the speed of light. The charge density ρ and the current density $\rho\mathbf{v}$ in the right-hand sides of (1.1) satisfy the *equation of continuity*

$$\frac{\partial \rho}{\partial t} + \text{div}\,(\rho\mathbf{v}) = 0 \qquad (1.3)$$

which, written in the integral form,

$$\frac{d}{dt}\int_{V_0} \rho\,d\tau = -\int_{\sigma_0} \rho v_n\,d\sigma \qquad (1.4)$$

expresses the fact that the change in the number of electrons inside the volume V_0 is equal to the number of electrons that pass through the surface σ_0 encompassing the volume. Let us put

$$\mathscr{E} = -\,\text{grad}\,\varphi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}, \qquad \mathscr{H} = \text{curl}\,\mathbf{A} \qquad (1.5)$$

where φ is the *scalar potential*, and $\mathbf{A}$ the *vector potential*. Both potentials we will subject to the usual condition (the *Lorentz condition*)

$$\text{div}\,\mathbf{A} + \frac{1}{c}\frac{\partial \varphi}{\partial t} = 0 \qquad (1.6)$$

By substituting (1.5) into the Maxwell equations (1.1) and (1.2) we find that Eqs. (1.2) are satisfied automatically and that Eqs. (1.1), if we use condition (1.6), read

$$\nabla^2\mathbf{A} - \frac{1}{c^2}\frac{\partial^2\mathbf{A}}{\partial t^2} = -\frac{4\pi}{c}\rho\mathbf{v}$$

$$\nabla^2\varphi - \frac{1}{c^2}\frac{\partial^2\varphi}{\partial t^2} = -4\pi\rho \qquad (1.7)$$

Let us assume that the space in which the fields are considered is unlimited. To make the solution of (1.7) unique we will make it a condition that there are no ingoing waves, that is, waves that arrive from infinity. This can be written

$$\lim_{r\to\infty} r\left(\frac{\partial f}{\partial r} + \frac{1}{c}\frac{\partial f}{\partial t}\right) = 0 \qquad (1.8)$$

where f is A_x, A_y, A_z or φ. The solution of (1.7) that satisfies this condition can be expressed in terms of *retarded potentials*, namely

$$\mathbf{A} = \int \left(\rho\frac{\mathbf{v}}{c}\right)_{t-|\mathbf{r}-\mathbf{r}'|/c} \frac{d\tau'}{|\mathbf{r}-\mathbf{r}'|}$$

$$\varphi = \int (\rho)_{t-|\mathbf{r}-\mathbf{r}'|/c} \frac{d\tau'}{|\mathbf{r}-\mathbf{r}'|} \qquad (1.9)$$

If ρ and ρv depend on time via a periodic factor, that is if

$$\rho = \rho_0 e^{i\omega t}, \qquad \rho v = (\rho v)_0 \, e^{i\omega t} \tag{1.10}$$

then the above formulas become

$$\mathbf{A} = e^{i\omega t} \frac{1}{c} \int (\rho \mathbf{v})_0 \, e^{-i\omega \mid r-r' \mid /c} \, \frac{d\tau'}{\mid r - r' \mid}$$

$$\varphi = e^{i\omega t} \int \rho_0 e^{-i\omega \mid r-r' \mid /c} \, \frac{d\tau'}{\mid r - r' \mid} \tag{1.11}$$

These are the classical formulas for the electromagnetic field generated by a continuous distribution of electric charges. We must now change them so as to account for the quantum nature of matter.

2. Charge density and current density

In the Maxwell equations (1.1) and (1.2) matter was characterized by charge density ρ and current density ρv. We must now find the quantum analogues to these quantities. The analogues can be of two types — mathematical expectations and operators. Since the classical quantities appear in our formulas as functions of time, we must choose a representation in which the operators will explicitly depend on time (see Section 14, Chapter III, Part I).

We start with the mathematical expectations. In the classical theory $\rho \, d\tau$ is the electric charge inside volume $d\tau$. In the quantum theory we find the mathematical expectation of a charge by multiplying the charge of one electron, $-e$, into the mathematical expectation of the number of electrons. If we have only one electron, the latter mathematical expectation is equal to the probability of the electron being in $d\tau$. This, as we know, is simply

$$\bar{\psi}\psi \, d\tau$$

The quantity ρ then corresponds to

$$\rho \to e\bar{\psi}\psi \tag{2.1}$$

Consider a small volume V_0 in which there are $N(V_0)$ electrons. The mathematical expectation of N is

$$\mathbf{M. \ E.} \ N(V_0) = \int_{V_0} \bar{\psi}\psi \, d\tau \tag{2.2}$$

We construct the time derivative of (2.2). Substituting for $\dot{\psi}$ and $\dot{\bar{\psi}}$ their expressions via the wave equation, we get

$$\frac{d}{dt} \int_{V_0} \bar{\psi}\psi \, d\tau = \frac{i}{\hbar} \int_{V_0} (\psi \overline{H\psi} - \bar{\psi}H\psi) \, d\tau \tag{2.3}$$

But according to the Schrödinger equation

$$H\psi = -\frac{\hbar^2}{2m}\nabla^2\psi + U\psi \qquad (2.4)$$

If we then use Eq. (2.4) and the fact that

$$\psi\nabla^2\bar{\psi} - \bar{\psi}\nabla^2\psi = \text{div}\,(\psi\,\text{grad}\,\bar{\psi} - \bar{\psi}\,\text{grad}\,\psi)$$

we can write formula (2.3) as

$$\frac{d}{dt}\int_{V_0}\bar{\psi}\psi\,d\tau = -\int_{V_0}\text{div}\,\mathbf{S}\,d\tau \qquad (2.5)$$

where vector $\mathbf{S}$ is defined as

$$\mathbf{S} = \frac{i\hbar}{2m}\,(\psi\,\text{grad}\,\bar{\psi} - \bar{\psi}\,\text{grad}\,\psi) \qquad (2.6)$$

Applying to (2.5) Gauss's integral theorem, we get

$$\frac{d}{dt}\int_{V_0}\bar{\psi}\psi\,d\tau = -\int_{\sigma_0}S_n\,d\sigma \qquad (2.7)$$

where n is the outer normal to the surface σ_0 that encompasses V_0. This can be interpreted as meaning that the change in the mathematical expectation of the number of electrons inside V_0 is equal to the mathematical expectation of the number of electrons passing through σ_0. The quantity $S_n\,d\sigma$ will then be the mathematical expectation of the number of electrons passing through elementary area $d\sigma$ in unit time. The numbers are added algebraically, that is, if $d\sigma$ is passed in both directions by equal numbers of electrons, $S_n = 0$. Vector $\mathbf{S}$ is therefore the *electron flux*, which implies that the classical current density has corresponding to it the vector quantity $-e\mathbf{S}$:

$$\rho\mathbf{v} \rightarrow -e\mathbf{S} \qquad (2.8)$$

Such a definition of $\mathbf{S}$ yields the equation

$$\frac{\partial\,(\bar{\psi}\psi)}{\partial t} + \text{div}\,\mathbf{S} = 0 \qquad (2.9)$$

from which it follows that the quantum analogues of ρ and $\rho\mathbf{v}$, (2.1) and (2.8), satisfy the equation of continuity (1.3).

We can transform expression (2.6) for $\mathbf{S}$ as follows. Put

$$\psi = ae^{i\alpha} \qquad (2.10)$$

where a and α are real. Then, and this is easy to verify,

$$\mathbf{S} = \frac{\hbar}{m}\,a^2\,\text{grad}\,\alpha \qquad (2.11)$$

which implies that S is parallel to the gradient of the phase of the wave function.

On the other hand, we can express the derivatives $\partial\psi/\partial x$, $\partial\psi/\partial y$, and $\partial\psi/\partial z$ in (2.6) in terms of the results of applying the operators p_x, p_y, and p_z. We then get

$$S_x = \frac{1}{2m}(\psi\overline{p_x\psi} + \bar{\psi}p_x\psi) \tag{2.12}$$

or

$$S_x = \frac{1}{2}(\psi\overline{\dot{x}\psi} + \bar{\psi}\dot{x}\psi) \tag{2.13}$$

and similar formulas for the other two components. Here $\dot{x}$, $\dot{y}$, and $\dot{z}$ are the operators of the respective components of velocity (see Section 1, Chapter I).

Formula (2.13) and the ones for the other two components are, from the formal aspect, a natural generalization of the classical quantities $\rho\dot{x}$, $\rho\dot{y}$, $\rho\dot{z}$ (divided by the electron charge, $-e$), since charge density ρ corresponds to $-e\bar{\psi}\psi$ and operators $\dot{x}$, $\dot{y}$, $\dot{z}$ to the components of velocity.

In Part V we will see that in Dirac's theory of the electron vector S can also be represented as (2.13) although the operators $\dot{x}$, $\dot{y}$, $\dot{z}$ are completely different in form from the ones in Schrödinger's theory.

Thus we have found the expressions for the mathematical expectations of the number of electrons in a given volume and the number of electrons passing through the surface encompassing this volume. We now turn to the problem of finding the operators.

If we have only one electron, the number of electrons (zero or unity) in volume V_0 is a function $f_{V_0}(x, y, z)$ of the position (x, y, z) of our electron, where f is determined as follows:

$$f_{V_0}(x, y, z) = 1 \quad \text{if } (x, y, z) \text{ lies inside } V_0$$
$$= 0 \quad \text{if } (x, y, z) \text{ lies outside } V_0 \tag{2.14}$$

For this reason in terms of x, y, z the operator $N(V_0)$ corresponding to the number of electrons will be multiplication into this function. The mathematical expectation of this operator will be exactly (2.2), which is what we expected. If we want to use the Heisenberg picture for operators, we must build the matrix with elements

$$(n \mid N \mid n') = \int \bar{\psi}_n(x, t) f \psi_{n'}(x, t) \, d\tau \tag{2.15}$$

where the $\psi_n(x, t)$ are a complete set of functions satisfying the Schrödinger equation, for example, the eigenfunctions of the

Hamiltonian. Since f is nonzero only inside V_0, where it is equal to unity, we get

$$(n \mid N \mid n') = \int_{V_0} \bar{\psi}_n \psi_{n'} \, d\tau \qquad (2.16)$$

and integration is carried out only over the points inside volume V_0.

After multiplying this matrix element, which is an element of Heisenberg's matrix for the operator $N(V_0)$, by the electron charge we come to the analogue of the classical quantity

$$\int_{V_0} \rho \, d\tau \rightarrow -e \int_{V_0} \bar{\psi}_n \psi_{n'} \, d\tau \qquad (2.17)$$

The charge density ρ then corresponds to

$$\rho_{nn'} = -e \bar{\psi}_n \psi_{n'} \qquad (2.18)$$

This comparison is done in the same way as in Section 8, Chapter I, where the elements of Heisenberg's matrix for coordinate x were compared with the classical expressions for the same quantity.

After we have found the quantum analogue of charge density ρ, we can derive the quantum analogue of the current density ρv in the same way that we derived (2.8) from (2.1).

Recalling (2.5) and (2.6), we get

$$\frac{d}{dt} \int_{V_0} \bar{\psi}_n \psi_{n'} \, d\tau = - \int_{V_0} \operatorname{div} \mathbf{S}_{nn'} \, d\tau \qquad (2.19)$$

where $\mathbf{S}_{nn'}$ is the vector

$$\mathbf{S}_{nn'} = \frac{i\hbar}{2m} (\psi_{n'} \operatorname{grad} \bar{\psi}_n - \bar{\psi}_n \operatorname{grad} \psi_{n'}) \qquad (2.20)$$

We can then relate ρv with

$$(\rho v)_{nn'} = -e \mathbf{S}_{nn'} \qquad (2.21)$$

Thus (2.12) and (2.13) transform into

$$(S_x)_{nn'} = \frac{1}{2m} (\psi_{n'} \overline{p_x \psi_n} + \bar{\psi}_n p_x \psi_{n'}) \qquad (2.22)$$

or

$$(S_x)_{nn'} = \frac{1}{2} (\psi_{n'} \overline{\dot{x} \psi_n} + \bar{\psi}_n \dot{x} \psi_{n'}) \qquad (2.23)$$

We obtained the quantum formulas for the charge and current densities assuming only one electron, but we can easily generalize this case for several electrons.

3. Frequencies and intensities

What remains is to substitute the elements of Heisenberg's matrices into the classical formulas of Section 1. First we note that if the $\psi_n(x, t)$ are the eigenfunctions of the Hamiltonian, that is

$$\psi_n(x, t) = e^{-iE_n t/\hbar}\psi_n^0(x) \tag{3.1}$$

then the dependence of $\rho_{nn'}$ and $(\rho v)_{nn'}$ on time is purely periodic with an angular frequency

$$\omega_{nn'} = 2\pi\nu_{nn'} = (E_n - E_{n'})/\hbar \tag{3.2}$$

The electromagnetic field (1.11) will obviously have the same frequency. Hence the *Bohr frequency relation*, according to which the frequency of radiation emitted by an atom is equal to the energy difference $E_n - E_{n'}$ between two stationary states divided by Planck's constant $h = 2\pi\hbar$. It is convenient then to link the emission process with the quantum transition ("jump") of the atom from one stationary state to another.

If we substitute (2.18) and (2.21) (the charge and current densities) into the formula for the potentials, (1.9) or (1.11), we have [2]

$$\mathbf{A}_{nn'} = -\frac{e}{c}\int \mathbf{S}_{nn'}e^{-i\omega_{nn'}|\mathbf{r}-\mathbf{r}'|/c}\frac{d\tau'}{|\mathbf{r}-\mathbf{r}'|}$$

$$\varphi_{nn'} = -e\int \bar{\psi}_n\psi_{n'}e^{-i\omega_{nn'}|\mathbf{r}-\mathbf{r}'|/c}\frac{d\tau'}{|\mathbf{r}-\mathbf{r}'|} \tag{3.3}$$

We will now assume that the emitted radiation has a wavelength

$$\lambda_{nn'} = 2\pi c/\omega_{nn'} \tag{3.4}$$

much greater than the dimensions of the atomic system. This implies that we can disregard the difference in the retardation of the electromagnetic fields emitted by different parts of the system. [3] On this assumption the factor $|\mathbf{r} - \mathbf{r}'|^{-1} e^{-i\omega_{nn'}|\mathbf{r}-\mathbf{r}'|/c}$ in the integrand will not change considerably in the region where $\bar{\psi}_n\psi_{n'}$ and $\mathbf{S}_{nn'}$ do not vanish, and in the expression for $\mathbf{A}_{nn'}$ we can disregard $\mathbf{r}'$ in comparison with $\mathbf{r}$ and take the factor outside the

---

[2] In the formulas e can be either the base of natural logarithms or the electron charge. But obviously there can be no confusion.

[3] This difference can be taken into account in the higher-order approximations. It is imperative to do so if the electromagnetic field vanishes (see the selection rules below).

integral sign. In the expression for $\varphi_{nn'}$ we can replace this factor by a linear function of position. In summary we get

$$\mathbf{A}_{nn'} = -\frac{e}{r} e^{-i\omega_{nn'}r/c} \int \mathbf{S}_{nn'} \, d\tau'$$

$$\varphi_{nn'} = -\frac{e}{r} e^{-i\omega_{nn'}r/c} \int \bar{\psi}_n \psi_{n'} \left(1 + \frac{i\omega_{nn'}}{c} \frac{\mathbf{r} \cdot \mathbf{r}'}{r}\right) d\tau' \qquad (3.5)$$

where according to (2.23)

$$\int \mathbf{S}_{nn'} \, d\tau = \frac{1}{2} \int (\psi_{n'} \overline{\mathbf{v}\psi_n} + \bar{\psi}_n \mathbf{v}\psi_{n'}) \, d\tau$$

Since the velocity operator is hermitian, both terms in the integral on the right-hand side are equal, which yields

$$\int \mathbf{S}_{nn'} \, d\tau = \int \bar{\psi}_n \mathbf{v}\psi_{n'} \, d\tau \qquad (3.6)$$

By virtue of the equations of motion the matrix element for velocity is equal to the time derivative of the corresponding matrix elements for position:

$$\int \bar{\psi}_n \mathbf{v}\psi_{n'} \, d\tau' = \frac{d}{dt} \int \bar{\psi}_n \mathbf{x}\psi_{n'} \, d\tau' = \frac{d\mathbf{x}_{nn'}}{dt} \qquad (3.7)$$

where

$$\mathbf{x}_{nn'} = \int \bar{\psi}_n \mathbf{x}\psi_{n'} \, d\tau' \qquad (3.8)$$

On the other hand, due to the orthonormality of the functions ψ_n we have

$$\int \bar{\psi}_n \psi_{n'} \, d\tau = \delta_{nn'} \qquad (3.9)$$

We then substitute these expressions into (3.5) and at first put $n = n'$. Since the diagonal element $\mathbf{x}_{nn'}$ does not depend on time, we find that

$$\mathbf{A}_{nn} = 0$$

$$\varphi_{nn} = -e/r \qquad (3.10)$$

which is the electrostatic field of an electron. On the other hand, if $n \neq n'$, we have

$$\mathbf{A}_{nn'} = -\frac{e}{cr} e^{-i\omega_{nn'}r/c} \frac{d\mathbf{x}_{nn'}}{dt}$$

$$\varphi_{nn'} = -\frac{ei\omega_{nn'}}{c} \frac{1}{r^2} e^{-i\omega_{nn'}r/c} (\mathbf{r} \cdot \mathbf{x}_{nn'}) \qquad (3.11)$$

In view of the fact that

$$\mathbf{x}_{nn'} = \mathbf{x}_{nn'}^0 e^{i\omega_{nn'}t}$$

we can write

$$\mathbf{A}_{nn'} = - \frac{ie\omega_{nn'}}{c} \mathbf{x}^0_{nn'} \frac{1}{r} e^{i\omega_{nn'}(t-r/c)} \tag{3.12}$$

instead of the first equation in (3.11).

We now use formulas (1.5) for the electric and magnetic fields. We assume that, although the wavelength is long on the atomic scale, it nevertheless is short compared to the distance to the point where the field is calculated. To simplify matters we will consider only the contribution of the vector potential. This is justified by the fact that the scalar potential contributes only to the numerical value of a constant in the final formulas. We will then have

$$\mathscr{E}_{nn'} = - \frac{e\omega^2_{nn'}}{c^2} \mathbf{x}^0_{nn'} \frac{1}{r} e^{i\omega_{nn'}(t-r/c)}$$

$$\mathscr{H}_{nn'} = \mathscr{E}_{nn'} \boldsymbol{\times} \operatorname{grad} r \tag{3.13}$$

These formulas provide an insight into the polarization and intensity of light of a frequency corresponding to a definite spectral line. If, for instance, for each pair (n, n') (which actually means a definite transition) one of the components of $\mathbf{x}_{nn'}$, say $z_{nn'}$, is nonzero and the two others, $x_{nn'}$ and $y_{nn'}$, are zero, the light is polarized along the z axis. But if $x_{nn'} \neq 0$, $y_{nn'} \neq 0$, and $z_{nn'} = 0$, the light is polarized in the xy plane. For some pairs (n, n') the matrix elements of all three coordinates, $x_{nn'}$, $y_{nn'}$ and $z_{nn'}$, might vanish. Then the lines corresponding to these pairs would be absent. Such transitions are said to be *forbidden*. In many cases there are rules that make it possible to decide whether a line is forbidden or not. These are called *selection rules*.

To estimate the radiation intensity we build the time average of the *Poynting vector*. For this we take the real parts $\mathscr{E}'$ and $\mathscr{H}'$ of $\mathscr{E}$ and $\mathscr{H}$. The Poynting vector is then

$$\mathbf{P} = \frac{c}{4\pi} (\mathscr{E} \boldsymbol{\times} \mathscr{H}) = \frac{c}{4\pi} \mathscr{E}'^2_{nn'} \operatorname{grad} r \tag{3.14}$$

and the time average of $\mathbf{P}$ is

$$\langle \mathbf{P} \rangle = \frac{e^2\omega^4_{nn'}}{8\pi c^3} |x^0_{nn'}|^2 \frac{1}{r^2} \operatorname{grad} r$$

$$= \frac{e^2\omega^4_{nn'}}{8\pi c^3} (|x^0_{nn'}|^2 + |y^0_{nn'}|^2 + |z^0_{nn'}|^2) \frac{1}{r^2} \operatorname{grad} r \tag{3.15}$$

Hence the intensity of the spectral line with frequency $\omega_{nn'}$ is proportional to

$$I_{nn'} = e^2\omega^4_{nn'} (|x^0_{nn'}|^2 + |y^0_{nn'}|^2 + |z^0_{nn'}|^2) \tag{3.16}$$

so that the ratio of intensities of two spectral lines is given by the ratio of the corresponding $I_{nn'}$'s.

If we introduce the *electric dipole moment* of the electron, **D**, in the following way:

$$D_x = -ex, \qquad D_y = -ey, \qquad D_z = -ez \qquad (3.17)$$

then formula (3.16) can be written as

$$I_{nn'} = \omega_{nn'}^4 \left[|(n|D_x|n')|^2 + |(n|D_y|n')|^2 + |(n|D_z|n')|^2 \right] \quad (3.18)$$

or simply

$$I_{nn'} = \omega_{nn'}^4 |(n|\mathbf{D}|n')|^2 \qquad (3.18^*)$$

In the case when the atom has several electrons, (3.18) can still serve as the measure of radiation intensity if **D** is understood to be the total electric moment of all the electrons. The selection rule will then state which elements of Heisenberg's matrix for total moment **D** are nonzero.

Let us illustrate the above facts by considering the one-dimensional oscillator. In Chapter I we found that the elements of Heisenberg's matrix for coordinate x are

$$x_{nn'} = e^{i\omega t} \left(\frac{n\hbar}{2m\omega} \right)^{1/2} \delta_{n-1\,n'} + e^{-i\omega t} \left(\frac{(n+1)\,\hbar}{2m\omega} \right)^{1/2} \delta_{n+1\,n'} \quad (3.19)$$

[see (8.8) of Chapter I]. It then follows that the nonzero matrix elements (and hence intensities) are those whose quantum numbers n and n' differ by unity:

$$n - n' = \pm 1 \qquad (3.20)$$

This is the selection rule for the oscillator. The frequency of these *allowed* transitions is equal to the oscillator's fundamental frequency, and there are no higher harmonics in the radiation. The measure of intensity for the oscillator is

$$I_{n-1\,n} = \frac{\hbar\omega^3 e^2}{2m}\, n \qquad (3.21)$$

Note that the intensity here is proportional to the quantum number, n.

4. Intensities in a continuous spectrum

For a continuous spectrum the formulas for intensities must be modified. Let us assume that an electron is ejected from the energy level E_n to infinity, where its kinetic energy is E (an *ab-*

sorption spectrum). In this case we speak not of the intensity of the monochromatic light of frequency ω, where

$$\omega = (E - E_n)/\hbar \qquad (4.1)$$

but of the intensity of light with a frequency range

$$\Delta\omega = \Delta E/\hbar \qquad (4.2)$$

This implies that in Heisenberg's matrices we must shift from the eigenfunction $\psi'_n(x)$ to the proper differential corresponding to the integral

$$\frac{1}{(\Delta E)^{1/2}} \Delta\Psi = \frac{1}{(\Delta E)^{1/2}} \int_E^{E+\Delta E} \psi(x, E) \, dE \qquad (4.3)$$

and normalized in a way such that

$$\lim_{\Delta E \to 0} \frac{1}{\Delta E} \int |\Delta\Psi|^2 \, d\tau = 1 \qquad (4.4)$$

In other respects our results are the same as for the discrete spectrum.

The above substitution forces us to replace the matrix element

$$x_{nn'} = (E_n \mid x \mid E_{n'}) \qquad (4.5)$$

with

$$(E_n \mid x \mid E)(\Delta E)^{1/2} = \frac{1}{(\Delta E)^{1/2}} \int \overline{\psi_n(x)} \, x \, \Delta\Psi \, d\tau \qquad (4.6)$$

But $\Delta\Psi$ of (4.6) is approximately

$$\Delta\Psi \approx \psi(x, E) \Delta E \qquad (4.7)$$

This approximation can only be used if substituting (4.7) for $\Delta\Psi$ does not change the convergence of the integral in $\Delta\Psi$. In our case it does not, since in (4.6) $\Delta\Psi$ is multiplied into $\overline{\psi_n(x)}$, which is a rapidly decreasing function.[4] For this reason we can write the matrix element (4.6) as

$$(E_n \mid x \mid E) = \int \overline{\psi_n(x)} \, x\psi(x, E) \, d\tau \qquad (4.8)$$

Now we use the same method for coordinates y and z and in (3.15) make the substitution

$$x_{nn'} \to (E_n \mid \mathbf{x} \mid E)(\Delta E)^{1/2} \qquad (4.9)$$

We get the intensity of light per interval $\Delta E = \hbar \, \Delta\omega$:

$$I_n(E) \Delta E = e^2\omega_n^4 \big[|(E_n \mid x \mid E)|^2 + |(E_n \mid y \mid E)|^2 + |(E_n \mid z \mid E)|^2 \big] \Delta E \qquad (4.10)$$

---

[4] In (4.4), however, we cannot use (4.7).

10*

We must use this formula instead of (3.16) if we are dealing with a continuous spectrum.

5. Perturbation of an atom by a light wave

When a plane monochromatic light wave impinges on an atomic system, it initiates additional radiation whose frequency can be equal (a) to that of the incident wave and (b) to the sum or difference of the incident-wave frequency and the natural frequencies of the atom (the *Raman effect*). The additional radiation interferes with the incident wave, which results in a plane wave of a different wavelength. The change in wavelength in a medium can be described by the medium's *index of refraction* or its *dielectric constant*. Both, generally speaking, will depend on the frequency. Hence there will be *dispersion of light*.

We will now briefly analyze the Raman effect. Our problem consists of two steps. First, we must find the approximate solution to the Schrödinger wave equation with a perturbation in the form of a light wave. Second, we must find the frequencies and intensities of the additional radiation and derive the formula for the dielectric constant.

We start with the perturbed wave equation. If the wavelength of the incident light is considerably longer than atomic dimensions, the perturbation Hamiltonian is approximately

$$U = -(D_x \mathscr{E}_x + D_y \mathscr{E}_y + D_z \mathscr{E}_z) = -(\mathbf{D} \cdot \mathscr{E}) \qquad (5.1)$$

where $\mathscr{E}$ is the electric field vector, and $\mathbf{D}$ is the vector of the electric dipole moment of the atom. When there is one valence electron,

$$D_x = -ex, \qquad D_y = -ey, \qquad D_z = -ez \qquad (5.2)$$

The wave is considered monochromatic, so that

$$\mathscr{E} = \frac{1}{2}\mathscr{E}_0 (e^{i\omega t} + e^{-i\omega t}) \qquad (5.3)$$

The time dependence of the perturbation Hamiltonian will then be

$$U = U_0 e^{i\omega t} + U_0^+ e^{-i\omega t} \qquad (5.4)$$

where U_0 and U_0^+ do not depend on time. The notation U_0^+ is justified by the fact that U_0^+ is the hermitian conjugate of U_0, in the sense of the linear operator theory.

If H^0 is the zeroth-order Hamiltonian, the wave equation whose approximate solution we are trying to find will have the following form:

$$H^0\psi + (U_0 e^{i\omega t} + U_0^+ e^{-i\omega t})\psi - i\hbar \frac{\partial \psi}{\partial t} = 0 \qquad (5.5)$$

We consider the external electric field to be weak, and in looking for ψ we limit ourselves to the first-order approximation. This implies that we must take ψ in form

$$\psi = \psi^* + v + w \qquad (5.6)$$

where v and w are small. If we put (5.6) into (5.5) and neglect terms of the order of Uv and Uw, we get

$$H^0\psi^* + H^0 v + H^0 w + U_0 e^{i\omega t}\psi^* + U_0^+ e^{-i\omega t}\psi^*$$
$$= i\hbar\frac{\partial \psi^*}{\partial t} + i\hbar\frac{\partial v}{\partial t} + i\hbar\frac{\partial w}{\partial t}$$

To satisfy this equation we will adjust the functions ψ^*, v, and w in a way such that

$$H^0\psi^* - i\hbar\frac{\partial \psi^*}{\partial t} = 0 \qquad (5.7)$$

$$H^0 v - i\hbar\frac{\partial v}{\partial t} = -U_0 e^{i\omega t}\psi^* \qquad (5.8)$$

$$H^0 w - i\hbar\frac{\partial w}{\partial t} = -U_0^+ e^{-i\omega t}\psi^* \qquad (5.8^*)$$

Equation (5.7) is the unperturbed wave equation. Its solution is

$$\psi^* = \psi_n^0 e^{-iE_n t/\hbar} \qquad (5.9)$$

If we then substitute (5.9) into Eqs. (5.8) and (5.8*), we see that these equations will be satisfied if

$$v = v_n^0 e^{i(\omega - E_n/\hbar)t} \qquad (5.10)$$

$$w = w_n^0 e^{-i(\omega + E_n/\hbar)t} \qquad (5.10^*)$$

where v_n^0 and w_n^0 do not depend on time. For the last two functions we have the following equations:

$$H^0 v_n^0 - (E_n - \hbar\omega) v_n^0 = -U_0 \psi_n^0 \qquad (5.11)$$

$$H^0 w_n^0 - (E_n + \hbar\omega) w_n^0 = -U_0^+ \psi_n^0 \qquad (5.11^*)$$

When solving these equations, we must assume that $E_n \pm \hbar\omega$ do not coincide with any of the eigenvalues of H^0, that is

$$E_n - E_{n'} \pm \hbar\omega \neq 0 \qquad (5.12)$$

or

$$|\omega_{nn'}| \neq |\omega| \qquad (5.13)$$

where

$$\omega_{nn'} = (E_n - E_{n'})/\hbar \qquad (5.14)$$

Condition (5.13) implies that there is no resonance between natural frequency $\omega_{nn'}$ of the atom and frequency ω of the light

wave. If there is resonance, our method of solving Eq. (5.5), based on the assumption that v_n^0 and w_n^0 are small, does not work.

When there is no resonance, that is, when condition (5.13) holds, Eqs. (5.11) and (5.11*) can be solved by the method of Section 2, Chapter II. If to simplify matters we assume that there are no degenerate eigenvalues and no continuous spectrum, we get

$$v_n^0 = - \sum_m \frac{(m \mid U_0 \mid n)}{E_m - E_n + \hbar\omega} \, \psi_m^0 \tag{5.15}$$

$$w_n^0 = - \sum_m \frac{(m \mid U_0^+ \mid n)}{E_m - E_n + \hbar\omega} \, \psi_m^0 \tag{5.15*}$$

where

$$(m \mid U_0 \mid n) = \int \bar{\psi}_m^0 U_0 \psi_n^0 \, d\tau \tag{5.16}$$

Since we assumed that the wavelength of the incident light is much longer than the atomic dimensions, the amplitude of $\mathscr{E}$ can be considered constant within the limits of integration in (5.16). Formulas (5.1), (5.3), and (5.4) then yield

$$(m \mid U \mid n) = - \frac{1}{2} \left[\mathscr{E}_x^0 (m \mid D_x^0 \mid n) + \mathscr{E}_y^0 (m \mid D_y^0 \mid n) + \mathscr{E}_z^0 (m \mid D_z^0 \mid n) \right]$$

$$= - \frac{1}{2} (m \mid \mathscr{E}^0 \cdot \mathbf{D}^0 \mid n) \tag{5.17}$$

where

$$(m \mid \mathbf{D}^0 \mid n) = \int \bar{\psi}_m^0 \mathbf{D} \psi_n^0 \, d\tau \tag{5.18}$$

Finally, using (5.6), (5.9), and (5.10), we can write the approximate solution to Eq. (5.5):

$$\psi_n = e^{-iE_n t/\hbar} \left(\psi_n^0 + v_n^0 e^{i\omega t} + w_n^0 e^{-i\omega t} \right) \tag{5.19}$$

where v_n^0 and w_n^0 are found by (5.15) and (5.15*).

6. The dispersion formula

In the classical electron theory we characterize the radiation of an atom by its electric dipole moment. The quantum analogue of this moment is, as we saw in Section 3, an element of a Heisenberg's matrix corresponding to the product of the electron charge by its coordinate, or the sum of such products if there are several electrons.

The incident wave induces an additional electric dipole moment, which for an optically homogeneous medium is proportional to the electric field. In the general case this moment will be a linear

vector function of the electric field components. It is the frequency dependence of the proportionality factor (or the factors of the vector function) that serves as the basis for explaining dispersion.

If we know the approximate solution ψ_n of the wave equation when the perturbation is in the form of a light wave, we can build the elements of Heisenberg's matrix for the electric dipole moment. Neglecting the squares and products of v_n^0 and w_n^0, we get

$$\int \bar{\psi}_n \mathbf{D}\psi_{n'} \, d\tau = e^{i\omega_{nn'}t} \left(n \,|\, \mathbf{D}^0 \,|\, n' \right)$$

$$+ \tfrac{1}{2} \mathbf{C}_{nn'} e^{i\,(\omega_{nn'}+\omega)\,t} + \tfrac{1}{2} \mathbf{C}_{nn'}^{+} e^{i\,(\omega_{nn'}-\omega)\,t} \tag{6.1}$$

where

$$\mathbf{C}_{nn'} = 2 \int \bar{w}_n^0 \mathbf{D}\psi_{n'}^0 \, d\tau + 2 \int \bar{\psi}_n^0 \mathbf{D}v_{n'}^0 \, d\tau \tag{6.2}$$

$$\mathbf{C}_{nn'}^{+} = 2 \int \bar{v}_n^0 \mathbf{D}\psi_{n'}^0 \, d\tau + 2 \int \bar{\psi}_n^0 \mathbf{D}w_{n'}^0 \, d\tau \tag{6.2*}$$

so that

$$\mathbf{C}_{nn'}^{+} = \overline{\mathbf{C}}_{n'n} \tag{6.3}$$

which justifies the "dagger" in (6.2*). If in (6.2) we now substitute (5.15) and (5.15*) for v_n^0 and w_n^0 and use (5.17), we get

$$\mathbf{C}_{nn'} = \sum_m \frac{(n \,|\, \mathscr{E}^0 \cdot \mathbf{D}^0 \,|\, m)\,(m \,|\, \mathbf{D}^0 \,|\, n')}{\hbar\,(\omega_{mn} - \omega)} + \sum_m \frac{(n \,|\, \mathbf{D}^0 \,|\, m)\,(m \,|\, \mathscr{E}^0 \cdot \mathbf{D}^0 \,|\, n')}{\hbar\,(\omega_{mn'} + \omega)} \tag{6.4}$$

We see that in addition to the natural frequency of the atom, $\omega_{nn'}$, formula (6.1) contains the sum and difference $\omega_{nn'} \pm \omega$ (the *Raman effect*). The terms

$$\mathbf{D}'_{nn'} = \tfrac{1}{2} \mathbf{C}_{nn'} e^{i\,(\omega_{nn'}+\omega)\,t} + \tfrac{1}{2} \mathbf{C}_{nn'}^{+} e^{i\,(\omega_{nn'}-\omega)\,t} \tag{6.5}$$

in (6.1) represent the additional dipole moment induced by the perturbation, the light wave. The diagonal element of this moment,

$$\mathbf{D}'_n = \mathbf{D}'_{nn} = \tfrac{1}{2} \mathbf{C}_{nn} e^{i\omega t} + \tfrac{1}{2} \mathbf{C}_{nn} e^{-i\omega t} \tag{6.6}$$

has the frequency of the incident wave. For this reason the diagonal element is the immediate counterpart of the classical additional moment.

Let us now turn to the dependence of $\mathbf{D}'_n$ on the electric field components. At $n = n'$ the expression (6.4) for $\mathbf{C}_{nn'}$ yields

$$(D'_x)_n = \mathrm{Re}\left[(\alpha_{xx})_n\,\mathscr{E}_x e^{i\omega t} + (\alpha_{xy})_n\,\mathscr{E}_y e^{i\omega t} + (\alpha_{xz})_n\,\mathscr{E}_z e^{i\omega t}\right]$$

$$(D'_y)_n = \mathrm{Re}\left[(\alpha_{yx})_n\,\mathscr{E}_x e^{i\omega t} + (\alpha_{yy})_n\,\mathscr{E}_y e^{i\omega t} + (\alpha_{yz})_n\,\mathscr{E}_z e^{i\omega t}\right]$$

$$(D'_z)_n = \mathrm{Re}\left[(\alpha_{zx})_n\,\mathscr{E}_x e^{i\omega t} + (\alpha_{zy})_n\,\mathscr{E}_y e^{i\omega t} + (\alpha_{zz})_n\,\mathscr{E}_z e^{i\omega t}\right] \tag{6.7}$$

where, for instance,

$$(\alpha_{xy})_n = \sum_m \left(\frac{(n|D_y^0|m)(m|D_x^0|n)}{\hbar(\omega_{mn} - \omega)} + \frac{(n|D_x^0|m)(m|D_y^0|n)}{\hbar(\omega_{mn} + \omega)} \right) \quad (6.8)$$

The other coefficients are expressed by similar formulas, in which x and y are changed to the appropriate labels.

The array

$$\begin{pmatrix} \alpha_{xx} & \alpha_{xy} & \alpha_{xz} \\ \alpha_{yx} & \alpha_{yy} & \alpha_{yz} \\ \alpha_{zx} & \alpha_{zy} & \alpha_{zz} \end{pmatrix} \quad (6.9)$$

(where we omitted label n) is a hermitian matrix, so that, for instance,

$$\alpha_{yx} = \bar{\alpha}_{xy} \quad (6.10)$$

Thus the additional dipole moment $\mathbf{D}'$ is a linear vector function of the electric field, and for complex coefficients in (6.9) the phase of $\mathbf{D}'$ does not coincide with the phase of $\mathscr{E}$. But if the α's are real (which is the case when the functions ψ_n^0 are real, for instance), the phases of $\mathbf{D}'$ and $\mathscr{E}$ coincide and Eqs. (6.7) read

$$(D_x')_n = (\alpha_{xx})_n \mathscr{E}_x + (\alpha_{xy})_n \mathscr{E}_y + (\alpha_{xz})_n \mathscr{E}_z$$
$$(D_y')_n = (\alpha_{yx})_n \mathscr{E}_x + (\alpha_{yy})_n \mathscr{E}_y + (\alpha_{yz})_n \mathscr{E}_z$$
$$(D_z')_n = (\alpha_{zx})_n \mathscr{E}_x + (\alpha_{zy})_n \mathscr{E}_y + (\alpha_{zz})_n \mathscr{E}_z \quad (6.11)$$

where $\mathscr{E}$ is the real electric field vector. A particularly interesting case is when

$$\alpha_{xx} = \alpha_{yy} = \alpha_{zz} = \alpha_n$$
$$\alpha_{yz} = \alpha_{zy} = \alpha_{xy} = 0 \quad (6.12)$$

and the dependence of $\mathbf{D}'$ on $\mathscr{E}$ simply reduces to proportionality:

$$\mathbf{D}_n' = \alpha_n \mathscr{E} \quad (6.13)$$

Up to now we have dealt with the additional dipole moment for one particle. To find the total electric dipole moment per unit volume, $\mathbf{D}^*$, we denote by N_r the number of particles in state r per unit volume and construct the sum

$$\mathbf{D}^* = \sum_r N_r \mathbf{D}_r' \quad (6.14)$$

which is

$$\mathbf{D}^* = \alpha \mathscr{E} \quad (6.15)$$

where

$$\alpha = \sum_r N_r \alpha_r \quad (6.16)$$

According to the classical electron theory the dielectric constant ε and the proportionality factor α in (6.15) are related in the following way:

$$\frac{3}{4\pi} \frac{\varepsilon - 1}{\varepsilon + 2} = \alpha \tag{6.17}$$

Hence our formulas connect the dielectric constant with atomic quantities.

We note that N_r of (6.16) depends on temperature. Classical Boltzmann statistics gives this dependence as

$$N_r = N \frac{\exp(-E_r/kT)}{\sum\limits_r \exp(-E_r/kT)} \tag{6.18}$$

where N is the total number of atoms per unit volume, and E_r is the energy of one atom in state r. As to α_r, it depends only on the properties of the particles and on the frequency of the incident wave, ω. This frequency dependence of α explains dispersion. We have just seen that ω enters the expression for α through the denominators $\omega_{mn} \pm \omega$, characteristic of the dispersion formula.

The theory presented in this section is no more than an outline, giving a general idea of the effect of light on atom. It is far from complete for several reasons. First, we have only briefly mentioned the relation between quantities referring to one atom or molecule (the dipole moment, for instance) and macroscopic quantities (for instance, the dielectric constant), and for the distribution by state we have confined ourselves to the classical formula (6.18). Second, even quantities referring to one atom were described semiclassically since we did not introduce the concept of light quanta, and schematically since we did not touch on the problem of degenerate eigenvalues, for instance, and did not clarify in what conditions (6.12) holds. Last, we said nothing about what happens if there is resonance and did not discuss the question of the width of a spectral line.

As we said at the beginning of this chapter, a full exposition of the theory of radiation goes beyond the subject of this book.

7. Penetration of a potential barrier by a particle

In this section we will consider a problem that in classical terms could be called the *penetration of a potential barrier* by a particle. In essence this problem is nonstationary, but we will use wave functions that are the eigenfunctions of the Hamiltonian as auxiliary quantities.

Let us assume that the potential U depends only on the distance from an attractive centre, that is, the problem is spherically

symmetric. We will also deal only with states in which the wave functions depend (apart from time) on r:

$$\psi = \psi(r, t) \tag{7.1}$$

(The general case is considered in Chapter IV.) If we put

$$\psi = \frac{f(r)}{r} \tag{7.2}$$

then the Schrödinger equation

$$-\frac{\hbar^2}{2m} \nabla^2 \psi + U\psi = i\hbar \frac{\partial \psi}{\partial t} \tag{7.3}$$

transforms into

$$-\frac{\hbar^2}{2m} \frac{\partial^2 f}{\partial r^2} + Uf = i\hbar \frac{\partial f}{\partial t} \tag{7.4}$$

For a state with definite energy (a stationary state) we can assume that

$$f(r, t) = f(r) e^{-iEt/\hbar} \tag{7.5}$$

Then Eq. (7.4) reads

$$-\frac{\hbar^2}{2m} \frac{d^2 f}{dr^2} + [U(r) - E]f = 0 \tag{7.6}$$

For the function ψ to remain finite everywhere, we must assume the following boundary condition:

$$f(0) = 0 \tag{7.7}$$

Now we make certain assumptions concerning the behaviour of $U(r)$. Let $U(r)$ first increase monotonically, starting from some definite value (or even from $-\infty$, which is the case for an electron in the Coulomb field of a nucleus, $U = -e^2/r$). In the process $U(r)$ attains a maximum value, after which it decreases without restriction. We will consider the values of E that are less than the maximum of $U(r)$. The difference $U(r) - E$ vanishes at two points, before and after the maximum, say at $r = r_1$ and $r = r_2$. Hence for r there are three distinct regions

$$\text{I. } 0 < r < r_1, \qquad \text{II. } r_1 < r < r_2, \qquad \text{III. } r_2 < r \tag{7.8}$$

How does f behave in each region? In the first, where $U < E$, function f will oscillate. If we put

$$S_1(r) = \int_0^r [2m(E - U)]^{1/2} \, dr \tag{7.9}$$

then in the semiclassical approximation (Section 15, Chapter III, Part I) and for r's that are smaller than r_1 (but not too close to r_1) we have

$$f = c \left(\frac{\partial^2 S_1}{\partial r \, \partial E} \right)^{1/2} \cos(S_1/\hbar + \alpha) \tag{7.10}$$

where c and α are constants. In the second region we put

$$S_2(r) = \int_{r_1}^{r} [2m(U - E)]^{1/2} \, dr \qquad (7.11)$$

Then approximately

$$f = \left(\frac{\partial^2 S_2}{\partial r \, \partial E}\right)^{1/2} (c_1 e^{S_2/\hbar} + c_2 e^{-S_2/\hbar}) \qquad (7.12)$$

In the third region

$$f = c' \left(\frac{\partial^2 S_3}{\partial r \, \partial E}\right)^{1/2} \cos(S_3/\hbar + \alpha') \qquad (7.13)$$

where

$$S_3 = \int_{r_2}^{r} [2m(E - U)]^{1/2} \, dr \qquad (7.14)$$

and c' and α' are new constants. In view of (7.2), function ψ differs from f by a factor, $1/r$, that tends to zero as $r \to \infty$. For this reason the behaviour of ψ in the first and third regions will be different and ψ will tend to zero as $r \to \infty$.

In the transition region, where the difference $U(r) - E$ in the Schrödinger equation (7.6) changes sign (that is, near the "turning" points $r = r_1$ and $r = r_2$), the above formulas for f and S do not hold; there the approximate solutions can be expressed using the *Airy functions*, which are the solutions of *Airy's differential equation*

$$w''(t) = tw(t) \qquad (7.15)$$

in which the coefficient of the unknown function passes through zero. We will not discuss the use of the Airy functions as applied to our problem.

The most interesting case is that of a very high barrier. Such a barrier is characterized by

$$\frac{1}{\hbar} S = \frac{1}{\hbar} S_2(r_2) \qquad (7.16)$$

being considerably greater than unity. The behaviour of f (and of ψ) will depend essentially on the coefficients in (7.12). If $c_1 = 0$, the first term in (7.12) vanishes and as r increases from r_1 to r_2 the function f rapidly decreases, so that the amplitude of ψ will be considerably smaller in region III than in region I. This means that the probability (more correctly, the probability density) of finding the particle inside the barrier is much greater than of finding it outside.

For arbitrary values of E we cannot make c_1 vanish. The condition $c_1(E) = 0$ selects certain values of E. (This was the case with the standard problem of Schrödinger's theory, where the

eigenvalues of the Hamiltonian were selected.) Let $E = E_0$ be
one of such values. The corresponding eigenfunction will be $\psi_{E_0}(r)$.

Since this function is nonzero outside the barrier, we cannot
consider it the correct function describing the particle inside the
barrier. To describe such a state we introduce a function $\psi_0(r)$
that is close to $\psi_{E_0}(r)$, namely

$$\psi_0(r) = \psi_{E_0}(r) \quad \text{at} \quad 0 < r < r_2$$

$$= 0 \qquad \text{at} \quad r_2 < r \qquad (7.17)$$

with normalization

$$\int_0^\infty |\psi_0(r)|^2 r^2\, dr = \int_0^{r_2} |\psi_{E_0}(r)|^2 r^2\, dr = 1 \qquad (7.18)$$

In the state described by $\psi_0(r)$ the particle is inside the barrier.
This state, however, is nonstationary since ψ_0 does not satisfy the
continuity condition at $r = r_2$, which is a necessary condition for
any energy eigenfunction. Nevertheless, the state will in a sense
be *quasi-stationary*. This means that if at time $t = 0$ it was de-
scribed by $\psi_0(r)$, for which the probability of finding the particle
inside the barrier is unity, in subsequent moments the probability
will be a slowly decreasing function of time. The law of such
slowly decreasing probability can be called the *law of decay of a
system in a quasi-stationary state*.

In classical mechanics the decay of a system would be as-
sociated with passage of particles over the potential barrier. If a
particle could always be localized in space, we would have to
say that in the region above the barrier the particle's kinetic
energy takes on negative values, which is impossible. On the
other hand, such phenomena as the escape of α-particles from an
atom's nucleus or the ionization of atoms in an external electric
field prove that decay of systems of such types is indeed possible.
If we now compare the two conclusions, we automatically deduce
that we cannot apply classical concepts to the above phenomena
and that the interpretation of such phenomena requires new
(quantum) concepts. But according to quantum mechanics we
cannot say that a particle is above a barrier until we show the
way this can be determined. In turn, to find the particle above the
barrier we must impart the lacking energy to it. This is the way
the above paradox is removed.

8. The law of decay of a quasi-stationary state

We can formulate the law of decay of a quasi-stationary state
of a system for a fairly general case if we introduce the *energy
distribution function* in this state.

Denote by x the set of all coordinates (or other variables, in terms of which the wave function is expressed). Let $\psi_0 = \psi(x, 0)$ be the initial value of the system's wave function, $\psi(x, t)$. We express ψ_0 as an integral of the eigenfunctions of the Hamiltonian, $\psi_E(x)$:

$$\psi(x, 0) = \int c(E) \psi_E(x) dE \qquad (8.1)$$

Then at $t > 0$ the state of the system will be

$$\psi(x, t) = \int e^{-iEt/\hbar} c(E) \psi_E(x) dE \qquad (8.2)$$

The probability $L(t)$ that at time t the system can be found in its initial state is the squared modulus of the scalar product

$$p(t) = \int \overline{\psi(x, 0)} \psi(x, t) dx \qquad (8.3)$$

so that we have

$$L(t) = |p(t)|^2 \qquad (8.4)$$

Using the completeness condition for the functions $\psi_E(x)$, we find that $p(t)$ can also be expressed as the scalar product of the coefficients of (8.1) and (8.2). We have

$$p(t) = \int e^{-iEt/\hbar} \overline{c(E)} c(E) dE \qquad (8.5)$$

But

$$dW(E) = w(E) dE = |c(E)|^2 dE \qquad (8.6)$$

is the energy distribution function for the initial state (which means that it is the distribution function at any time $t > 0$). Hence (8.5) reads

$$p(t) = \int e^{-iEt/\hbar} w(E) dE = \int e^{-iEt/\hbar} dW(E) \qquad (8.7)$$

Thus the probability that the system has not yet decayed at time t is

$$L(t) = |p(t)|^2 = \left| \int e^{-iEt/\hbar} dW(E) \right|^2 \qquad (8.8)$$

which implies that

The law of decay of state ψ_0 depends only on the energy distribution function.

With the appropriate choice of integral distribution function $W(E)$, Eq. (8.8) is valid for a discontinuous $W(E)$ (discrete spectrum).

We note that the law of decay can be the same for two different states provided their energy distribution functions are the same. We must also note that time t in the expression for the probability of decay is counted off starting from the (latest) moment when

we can still say that the atom (or system) has not yet decayed; the state of the undecayed atom, ψ_0, does not change. If we wish, we can say that the atom does not age but decays suddenly. This corollary is valid for any law of decay, not only the exponential law.

Interestingly, in formula (8.7) it is not the probability amplitude (wave function) that is subject to a Fourier transformation, which would be the case in quantum mechanics, but the probability itself. According to the nomenclature of probability theory $p(t)$ is the characteristic function for E.

The properties of the Fourier integral yield the relation between the rate of decay and the smoothness of the distribution function. Let us first investigate the conditions in which decay is possible at all.

If there is a *differential energy-distribution function* (probability density) $w(E)$, the *integral distribution function* $W(E)$ is related to it as

$$W(E') - W(E) = \int_E^{E'} w(E)\, dE \qquad (8.9)$$

where E' and E are any two values of energy. If $E' > E$, the above expression is obviously the probability that the energy of the system lies between E and E'. If there is no discrete spectrum, $W(E)$ is continuous for any initial state. But if there is a discrete spectrum, $W(E)$ will be continuous only if in the initial state all probabilities concerning the discrete spectrum vanish.

Assume that $W(E)$ is continuous. Then (8.9) implies that the continuity of $W(E)$ is equivalent to the absolute integrability of $w(E)$ (in the usual meaning). But if $w(E)$ is absolute integrable, the value of (8.5) tends to zero as $t \to \infty$.

Thus the continuity of $W(E)$ yields

$$L(t) \to 0 \quad \text{as} \quad t \to \infty \qquad (8.10)$$

where according to (8.7) and (8.8), $L(t)$ is the probability that at time t the system has not yet decayed. On the other hand we can show that the continuity of $W(E)$ follows from (8.10). We come to the conclusion that

A system decays if and only if the integral energy-distribution function is continuous.

In many problems the energy distribution function satisfies much stricter conditions than the simple continuity of $W(E)$. For instance, when a particle being in a potential well penetrates a bar-

rier, the probability density $w(E)$ will be a meromorphic function [5] of a complex variable (this problem was considered in Section 7). Since at real values of E the function $w(E)$ is real, its poles lie symmetrically with respect to the real axis, and the residues at each symmetric pair of poles will be complex conjugates (the physical meaning of $w(E)$ as probability density implies that it cannot have any poles on the real axis). Let the pair of poles closest to the real axis be

$$E = E_0 \pm i\Gamma, \qquad \Gamma > 0 \qquad (8.11)$$

Let the next pair of poles have imaginary parts $\pm i\Gamma'$. It is easy to see that if t is so great that

$$\left| \frac{\Gamma' - \Gamma}{\hbar} t \right| \gg 1 \qquad (8.12)$$

then the value of integral (8.7) will be determined by the residues at the two poles (8.11), whereas none of the other poles will contribute. So we can assume that $w(E)$ has only two poles, which means that we can put [6]

$$w(E) = \frac{1}{\pi} \frac{\Gamma}{(E - E_0)^2 + \Gamma^2} \qquad (8.13)$$

This is the dispersion formula for the energy distribution.

If we now substitute (8.13) into (8.7), we get

$$p(t) = e^{-iEt/\hbar - \Gamma |t|/\hbar} \qquad (8.14)$$

and hence

$$L(t) = e^{-2\Gamma |t|/\hbar} \qquad (8.15)$$

Thus to obtain the usual exponential form of the decay law it is sufficient to assume that $w(E)$ is meromorphic — an assumption that can be proved by analyzing the Schrödinger equation of the problem.

---

[5] A function is meromorphic throughout a region D if and only if its only singularities throughout D are poles.

[6] We assume that $\Gamma \ll E_0 - E^*$, where E^* is the lower limit of integration in (8.7); usually $E^* = 0$.

Chapter IV

AN ELECTRON IN A CENTRAL FIELD

1. General remarks

The problem of describing the states of an electron in a central field is of great practical importance, since its solution gives us not only the theory of the spectrum of hydrogen (an electron moving in the *Coulomb field*) but the approximate theory of the spectra of atoms with one valence electron (the sodium atom, for instance).

In the *hydrogen atom* the electron is in the electrostatic Coulomb field of the nucleus, so that the potential energy (or simply potential) is

$$U(r) = -\frac{e^2}{r}$$

In atoms with several electrons the electrons lose their individuality in a way, and, generally speaking, we cannot consider the states of separate electrons and describe them by using wave functions ψ that depend on the coordinates of each electron respectively. Instead we must examine the state of the atom as a whole and describe it by a wave function that depends on the coordinates of all the electrons. We must also take into account the intrinsic degree of freedom (the so-called spin) and the symmetry of the wave function with respect to the interchange of any two electrons (exchange symmetry). The many-electron problem is dealt with in Part IV. We will only remark that in a certain approximation the wave function of the atom as a whole can be expressed in terms of the wave function of each separate electron. In this case for the "one-electron" wave functions we get equations similar to those of the one-body problem (with certain additional terms). In view of this for an atom with one valence electron we can build an equation for the electron wave function and speak of the electron being in the field of the nucleus and the other (the inner) electrons. As in the case of the hydrogen atom, this field will be spherically symmetric, but it will not be a Coulomb field. Considering what has been said, the case involving the non-Coulomb field $U(r)$, which depends only on the distance from the nucleus, is of great importance to physics.

Schrödinger's theory gives a true picture in general outline of atoms with one valence electron. Only a few details, notably, the

fine structure of energy levels (the existence of doublets), cannot be derived from the Schrödinger equation but can be interpreted by Dirac's theory, which takes the theory of relativity into account. More than that, Dirac's theory is needed to explain the behaviour of an atom in a magnetic field (the Zeeman effect). True, the Schrödinger equation can be generalized for the case of a magnetic field, but since the corrections due to the magnetic field and the theory of relativity are of the same order of magnitude, they must be considered simultaneously. Dirac's theory of the electron is elaborated in Part V.

2. Conservation of angular momentum [7]

The Schrödinger wave equation for an electron in a field with potential energy $U(x, y, z)$ is

$$H\psi - i\hbar \frac{\partial \psi}{\partial t} = 0 \qquad (2.1)$$

where the Hamiltonian is

$$H = \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2) + U(x, y, z) \qquad (2.2)$$

We assume that U depends only on the distance from the atom's nucleus, the latter being fixed and at the origin of a coordinate system:

$$U(x, y, z) = U(r), \qquad r = (x^2 + y^2 + z^2)^{1/2} \qquad (2.3)$$

The wave equation will then be

$$\frac{1}{2m}(p_x^2 + p_y^2 + p_z^2)\psi + U(r)\psi - i\hbar \frac{\partial \psi}{\partial t} = 0 \qquad (2.4)$$

or, if we express p_x, p_y, p_z in terms of position derivatives,

$$-\frac{\hbar^2}{2m}\nabla^2\psi + U(r)\psi - i\hbar \frac{\partial \psi}{\partial t} = 0 \qquad (2.5)$$

where ∇^2 is the Laplacian operator.

In the case of a central field in classical mechanics we have the following conservation law:

The components of angular momentum about the origin of coordinates

$$m_x = yp_z - zp_y$$
$$m_y = zp_x - xp_z$$
$$m_z = xp_y - yp_x \qquad (2.6)$$

are constants of the motion.

---

[7] The law of conservation of angular momentum for a particle moving in a central field is sometimes called the *area integral.*

But will these quantities be constants of the motion in quantum mechanics if we assume them to be the operators considered in Section 7, Chapter III, Part I? For this we must see whether they commute with the Hamiltonian. One way is to use the Poisson bracket of Section 7, Chapter III, Part I. However, it is simpler to find the commutation relation directly.

We start with

$$m_z \psi = \frac{\hbar}{i} \left(x \frac{\partial \psi}{\partial y} - y \frac{\partial \psi}{\partial x} \right)$$

Hence

$$(\nabla^2 m_z) \psi - m_z \nabla^2 \psi = \frac{\hbar}{i} \left[\nabla^2 \left(x \frac{\partial \psi}{\partial y} - y \frac{\partial \psi}{\partial x} \right) - x \frac{\partial \nabla^2 \psi}{\partial y} + y \frac{\partial \nabla^2 \psi}{\partial x} \right]$$

$$= \frac{\hbar}{i} \left(2 \frac{\partial^2 \psi}{\partial x \, \partial y} - 2 \frac{\partial^2 \psi}{\partial x \, \partial y} \right) = 0$$

We can prove in a similar way that m_z commutes with $U(r)$, namely

$$(U m_z) \psi - m_z U \psi = \frac{\hbar}{i} \left[U \left(x \frac{\partial \psi}{\partial y} - y \frac{\partial \psi}{\partial x} \right) - x \frac{\partial U \, \psi}{\partial y} + y \frac{\partial U \, \psi}{\partial x} \right]$$

$$= \frac{\hbar}{i} \left(y \frac{\partial U}{\partial x} - x \frac{\partial U}{\partial y} \right) \psi = 0$$

since if $U = U(r)$ then

$$y \frac{\partial U}{\partial x} - x \frac{\partial U}{\partial y} = 0$$

Thus m_z commutes both with the Laplacian operator and with the potential energy, which implies that it commutes with the Hamiltonian. In view of the symmetry between the coordinates x, y, z the same is true for m_x and m_y. Hence

$$H m_x - m_x H = 0$$
$$H m_y - m_y H = 0$$
$$H m_z - m_z H = 0 \qquad (2.7)$$

which means that

The components of angular momentum are constants of the motion in quantum mechanics.

However, these operators do not commute with each other. Indeed, we already know that

$$m_y m_z - m_z m_y = i\hbar m_x$$
$$m_z m_x - m_x m_z = i\hbar m_y$$
$$m_x m_y - m_y m_x = i\hbar m_z \qquad (2.8)$$

which implies that the physical quantities m_x, m_y, m_z cannot have definite values simultaneously (with the exception of zero values). Let us show that

$$\mathbf{m}^2 = m_x^2 + m_y^2 + m_z^2 \tag{2.9}$$

is an operator that commutes with each of the operators m_x, m_y, m_z. We will call this new operator the *square of angular momentum*. We have

$$m_x^2 m_z - m_z m_x^2 = m_x (m_x m_z - m_z m_x) + (m_x m_z - m_z m_x) m_x$$
$$= -i\hbar (m_x m_y + m_y m_x)$$
$$m_y^2 m_z - m_z m_y^2 = i\hbar (m_x m_y + m_y m_x)$$
$$m_z^2 m_z - m_z m_z^2 = 0$$

Adding up the left- and right-hand sides, we get

$$\mathbf{m}^2 m_z - m_z \mathbf{m}^2 = 0$$

which in view of the symmetry between x, y, z is actually three commutation relations:

$$\mathbf{m}^2 m_x - m_x \mathbf{m}^2 = 0$$
$$\mathbf{m}^2 m_y - m_y \mathbf{m}^2 = 0$$
$$\mathbf{m}^2 m_z - m_z \mathbf{m}^2 = 0 \tag{2.10}$$

Equations (2.10) express the physical fact that the square of angular momentum and any of the components of angular momentum can simultaneously have definite values.

On the other hand, since each of the operators m_x, m_y, m_z commutes with the Hamiltonian, the sum of their squares also commutes with it. Hence $\mathbf{m}^2$ will be a constant of the motion:

$$H\mathbf{m}^2 - \mathbf{m}^2 H = 0 \tag{2.11}$$

We note that H is also a constant of the motion. Thus we have three operators, m_z (say), $\mathbf{m}^2$, and H, which commute with each other and are constants of the motion. The general theory then states that we can choose the function that satisfies Eq. (2.1) so that it is simultaneously the eigenfunction of all three operators and hence is a solution to the following equations:

$$H\psi = E\psi \tag{2.12}$$

$$\mathbf{m}^2\psi = \lambda\psi \tag{2.12*}$$

$$m_z\psi = m_z'\psi \tag{2.12**}$$

3. Operators in spherical coordinates.
Separation of variables

Since in the problem under consideration the field is spherically symmetric, in studying the operators in $(2.12)-(2.12^{**})$ it is convenient to introduce spherical coordinates r, θ, φ, which are related to Cartesian coordinates as

$$x = r \sin\theta \cos\varphi, \qquad y = r\sin\theta\sin\varphi, \qquad z = r\cos\theta \qquad (3.1)$$

Let us express the operators m_x, m_y, m_z in terms of θ and φ:

$$m_x\psi = \frac{\hbar}{i}\left(y\frac{\partial\psi}{\partial z} - z\frac{\partial\psi}{\partial y}\right) = i\hbar\left(\sin\varphi\frac{\partial\psi}{\partial\theta} + \cot\theta\cos\varphi\frac{\partial\psi}{\partial\varphi}\right)$$

$$m_y\psi = \frac{\hbar}{i}\left(z\frac{\partial\psi}{\partial x} - x\frac{\partial\psi}{\partial z}\right) = i\hbar\left(-\cos\varphi\frac{\partial\psi}{\partial\theta} + \cot\theta\sin\varphi\frac{\partial\psi}{\partial\varphi}\right)$$

$$m_z\psi = \frac{\hbar}{i}\left(x\frac{\partial\psi}{\partial y} - y\frac{\partial\psi}{\partial x}\right) = -i\hbar\frac{\partial\psi}{\partial\varphi} \qquad (3.2)$$

Then $\mathbf{m}^2 = m_x^2 + m_y^2 + m_z^2$ will operate on ψ in the following way:

$$\mathbf{m}^2\psi = -\hbar^2\left(\sin\varphi\frac{\partial}{\partial\theta} + \frac{\cos\varphi}{\tan\theta}\frac{\partial}{\partial\varphi}\right)\left(\sin\varphi\frac{\partial\psi}{\partial\theta} + \frac{\cos\varphi}{\tan\theta}\frac{\partial\psi}{\partial\varphi}\right)$$

$$-\hbar^2\left(\cos\varphi\frac{\partial}{\partial\theta} - \frac{\sin\varphi}{\tan\theta}\frac{\partial}{\partial\varphi}\right)\left(\cos\varphi\frac{\partial\psi}{\partial\theta} - \frac{\sin\varphi}{\tan\theta}\frac{\partial\psi}{\partial\varphi}\right)$$

$$-\hbar^2\frac{\partial^2\psi}{\partial\varphi^2}$$

or simply

$$\mathbf{m}^2\psi = -\hbar^2\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2\psi}{\partial\varphi^2}\right] \qquad (3.3)$$

What we have obtained is the differential operator that appears in potential theory in the study of the equation for *spherical harmonics* (or *functions*) $Y_l(\theta, \varphi)$:

$$\frac{1}{\sin^2\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Y_l}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2 Y_l}{\partial\varphi^2} + l(l+1)Y_l = 0 \qquad (3.4)$$

where the integer l $(l = 0, 1, 2, \ldots)$ is the *degree* of the spherical harmonic. If we compare (3.4) with the eigenvalue equation for $\mathbf{m}^2$, (2.12^*), we conclude that the eigenvalues of $\mathbf{m}^2$ are

$$\lambda = \hbar^2 l(l+1), \qquad l = 0, 1, 2, \ldots \qquad (3.5)$$

To find the new form of H we use the well-known expression for the Laplacian operator in spherical coordinates. This yields

$$H\psi = -\frac{\hbar^2}{2m}\left\{\frac{\partial^2\psi}{\partial r^2} + \frac{2}{r}\frac{\partial\psi}{\partial r}\right.$$

$$\left. + \frac{1}{r^2}\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2\psi}{\partial\varphi^2}\right]\right\}$$

$$+ U(r)\psi \qquad (3.6)$$

We see that the right-hand side has the same derivatives with respect to θ and φ as $\mathbf{m}^2$. As to the derivatives with respect to r, we can write them as

$$- \hbar^2 \left(\frac{\partial^2 \psi}{\partial r^2} + \frac{2}{r} \frac{\partial \psi}{\partial r} \right) = p_r^{*2} \psi \tag{3.7}$$

where p_r^* is an operator:

$$p_r^* \psi = \frac{\hbar}{ir} \frac{\partial}{\partial r} (r\psi) = \frac{\hbar}{i} \left(\frac{\partial \psi}{\partial r} + \frac{\psi}{r} \right) \tag{3.8}$$

which we can interpret, by analogy with $p_x = -i\hbar(\partial/\partial x)$, as the operator of the radial component of angular momentum. If we use $\mathbf{m}^2$ and p_r^{*2} in H, we get

$$H = \frac{1}{2m} \left(p_r^{*2} + \frac{1}{r^2} \mathbf{m}^2 \right) + U(r) \tag{3.9}$$

This energy operator corresponds in form to the classical Hamiltonian function in spherical coordinates.

The eigenvalue equations for the simultaneous eigenfunctions of H and $\mathbf{m}^2$ can then be written in the following form:

$$H\psi = \frac{1}{2m} \left(p_r^{*2} + \frac{1}{r^2} \mathbf{m}^2 \right) \psi + U(r) \psi = E\psi \tag{3.10}$$

$$\mathbf{m}^2 \psi = - \hbar^2 \left[\frac{1}{\sin\theta} \frac{\partial}{\partial\theta} \left(\sin\theta \frac{\partial\psi}{\partial\theta} \right) + \frac{1}{\sin^2\theta} \frac{\partial^2\psi}{\partial\varphi^2} \right] = \hbar^2 l(l+1) \psi \tag{3.11}$$

By using (3.8) and (3.11) we can express (3.10) as

$$- \frac{\hbar^2}{2m} \left[\frac{\partial^2\psi}{\partial r^2} + \frac{2}{r} \frac{\partial\psi}{\partial r} - \frac{l(l+1)}{r^2} \psi \right] + U(r) \psi = E\psi \tag{3.12}$$

We note that Eq. (3.11) contains explicitly only the variables θ and φ, and Eq. (3.12) the variable r. For this reason we can look for the solution of these equations in the form of the product of a function depending only on r and a function depending only on θ and φ. More than that, since ψ must satisfy the wave equation

$$H\psi = E\psi = i\hbar \frac{\partial\psi}{\partial t} \tag{3.13}$$

we can introduce the exponential factor $e^{-iEt/\hbar}$ and thus put

$$\psi = e^{-iEt/\hbar} \psi^0 (r, \theta, \varphi) \tag{3.14}$$

where, as we have just said,

$$\psi^0 (r, \theta, \varphi) = R(r) Y_l (\theta, \varphi) \tag{3.15}$$

The function that depends on angles θ and φ we have interpreted as a spherical harmonic of degree l since it satisfies Eq. (3.4), which coincides with (3.11). The *radial function*, $R(r)$, must satisfy

Eq. (3.12), which we write in the form

$$\frac{d^2R}{dr^2} + \frac{2}{r}\frac{dR}{dr} - \frac{l(l+1)}{r^2}R + \frac{2m}{\hbar^2}[E - U(r)]R = 0 \quad (3.16)$$

Thus knowing the laws of conservation of energy and of angular momentum enables us to separate the variables, that is, to reduce the solution of the wave equation in four variables (t, r, θ, φ) to the solution of simpler equations in a smaller number of variables.

4. Solution of the differential equation for spherical harmonics

We have seen that the eigenvalue equation (3.11) for the eigenfunctions of the square of angular momentum coincides with Eq. (3.4) for spherical harmonics. Hence studying angular momentum reduces to studying spherical harmonics.

Let us find the simultaneous eigenfunctions of operators m_z and $\mathbf{m}^2$. The spherical harmonic $Y_l(\theta, \varphi)$ will be an eigenfunction of m_z if it satisfies the equation

$$-i\hbar\frac{\partial Y_l}{\partial\varphi} = m_z'Y_l \quad (4.1)$$

whose solution is

$$Y_l(\theta, \varphi) = \Theta(\theta)e^{im_z'\varphi/\hbar}$$

To ensure that Y_l is a single-valued function of position in space we must see that it is a periodic function in φ with period 2π. This in turn means that the eigenvalues m_z' are equal to

$$m_z' = m\hbar, \qquad m = 0, \pm 1, \pm 2, \ldots \quad (4.2)$$

(Incidentally, we came to the same result in Section 7, Chapter III, Part I.) Hence

$$Y_l(\theta, \varphi) = \Theta(\theta)e^{im\varphi} \quad (4.3)$$

Substituting into Eq. (3.4), we get

$$\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{d\Theta}{d\theta}\right) - \frac{m^2}{\sin^2\theta}\Theta + l(l+1)\Theta = 0 \quad (4.4)$$

If we introduce a new variable

$$x = \cos\theta \quad (4.5)$$

then Eq. (4.4) reads

$$\frac{d}{dx}\left[(1-x^2)\frac{d\Theta}{dx}\right] - \frac{m^2}{1-x^2}\Theta + l(l+1)\Theta = 0 \quad (4.6)$$

The singular points of this equation are $x = \pm 1$. Indeed, if we solve it for the second derivative, the coefficients will become

infinitely great at $x = \pm 1$. Next, if we consider l to be an undefined parameter, we can show that Eq. (4.6) has a solution that at $x = \pm 1$ remains finite only if l is an integer. This implies that spherical harmonics are the only solutions of (4.6) that satisfy the stated conditions, that is, the only eigenfunctions of $\mathbf{m}^2$. Let us find the solution of (4.6) for integral l's.

First we consider the case $m = 0$. Put

$$y = (x^2 - 1)^l$$

and take the logarithmic derivative with respect to y:

$$\frac{y'}{y} = \frac{2lx}{x^2 - 1}$$

or

$$(1 - x^2)\frac{dy}{dx} + 2lxy = 0$$

We differentiate the last equation $k + 1$ times with respect to x and put

$$z = \frac{d^k y}{dx^k} = \frac{d^k}{dx^k}(x^2 - 1)^l \tag{4.7}$$

This yields

$$(1 - x^2)\frac{d^2 z}{dx^2} - (2k - 2l + 2)x\frac{dz}{dx} + (2l - k)(k + 1)z = 0 \tag{4.8}$$

If we put $k = l$, we get an equation that at $m = 0$ coincides with (4.6). The solution of this equation that at $x = 1$ is unity is denoted $P_l(x)$ and is called the *Legendre polynomial of degree l*. It differs from (4.7) at $k = l$ only by a constant factor. If we determine this factor from $P_l(1) = 1$, we get the *Rodrigues formula*

$$P_l(x) = \frac{1}{2^l l!}\frac{d^l}{dx^l}(x^2 - 1)^l \tag{4.9}$$

Hence the polynomial satisfies the equation

$$\frac{d}{dx}\left[(1 - x^2)\frac{dP_l}{dx}\right] + l(l + 1)P_l = 0 \tag{4.10}$$

which is a particular case of (4.6).

Let us now consider the general case $m \neq 0$. We introduce the substitution

$$\Theta = (1 - x^2)^{m/2}v \tag{4.11}$$

which gives the following equation for v:

$$(1 - x^2)\frac{d^2 v}{dx^2} - (2m + 2)x\frac{dv}{dx} + (l - m)(l + m + 1)v = 0 \tag{4.12}$$

If we had put

$$\Theta = (1 - x^2)^{-m/2} w \qquad (4.13)$$

instead of (4.11), we would have had an equation for w that differs from (4.12) only in the sign of m, namely

$$(1 - x^2)\frac{d^2w}{dx^2} + (2m - 2) x \frac{dw}{dx} + (l + m)(l - m + 1) w = 0 \qquad (4.14)$$

Both equations, (4.12) and (4.14), are of the type of Eq. (4.8), and for (4.12) integer k is $l + m$, whereas for (4.14) it is $l - m$. For this reason we must put

$$v = c_1 \frac{d^{l+m}}{dx^{l+m}} (x^2 - 1)^l \qquad (4.15)$$

$$w = c_2 \frac{d^{l-m}}{dx^{l-m}} (x^2 - 1)^l \qquad (4.16)$$

By equating the two substitutions for Θ, (4.11) and (4.13), we get

$$\Theta = c_1 (1 - x^2)^{m/2} \frac{d^{l+m}}{dx^{l+m}} (x^2 - 1)^l$$

$$= c_2 (1 - x^2)^{-m/2} \frac{d^{l-m}}{dx^{l-m}} (x^2 - 1)^l \qquad (4.17)$$

To find the ratio of the two constants, c_1/c_2, it is sufficient to equate (4.11) with (4.13) for any particular value of x. Computations yield

$$c_1 (l + m)! = c_2 (- 1)^m (l - m)! \qquad (4.18)$$

Usually c_1 is chosen such that

$$c_1 = \frac{1}{2^l l!} \qquad (4.19)$$

which gives

$$c_2 = \frac{1}{2^l l!} (- 1)^m \frac{(l + m)!}{(l - m)!} \qquad (4.20)$$

Also, the corresponding solution to Eq. (4.6) is denoted by $P_l^m(x)$. Thus

$$P_l^m (x) = (1 - x^2)^{m/2} \frac{d^{l+m}}{dx^{l+m}} \frac{(x^2 - 1)^l}{2^l l!} \qquad (4.21)$$

which is also

$$P_l^m (x) = (- 1)^m \frac{(l + m)!}{(l - m)!} (1 - x^2)^{-m/2} \frac{d^{l-m}}{dx^{l-m}} \frac{(x^2 - 1)^l}{2^l l!} \qquad (4.22)$$

These functions satisfy the equation

$$\frac{d}{dx}\left[(1-x^2)\frac{dP_l^m}{dx}\right] - \frac{m^2}{1-x^2}P_l^m + l(l+1)P_l^m = 0 \quad (4.23)$$

and represent the solutions that at $x = \pm 1$ are finite. (They are sometimes called the *associated Legendre polynomials of the first kind*.)

Formulas (4.21) and (4.22) give $P_l^m(x)$ for both positive and negative values of the integer m, and from comparing (4.21) with (4.22) it follows that

$$P_l^{-m}(x) = (-1)^m \frac{(l-m)!}{(l+m)!} P_l^m(x) \quad (4.24)$$

Also, (4.21) and (4.22) at $|m| > l$ vanish, which implies that for this case there are no solutions of (4.6) that at $x = \pm 1$ remain finite. Hence m (for a fixed l) can take on only the values

$$m = -l, \ -l+1, \ldots, l-1, \ l \quad (4.25)$$

altogether $2l + 1$ values. The inequality

$$|m| \leqslant l \quad (4.26)$$

stems from the physical meaning of these quantities. Indeed, to within a factor $\hbar^2$, the quantity m^2 is an eigenvalue of operator m_z^2, and $l(l+1)$ is an eigenvalue of $\mathbf{m}^2 = m_x^2 + m_y^2 + m_z^2$. This results in that [8]

$$m^2 \leqslant l(l+1) < \left(l + \frac{1}{2}\right)^2$$

which yields

$$|m| < l + \frac{1}{2}$$

Since $|m|$ and l are integers, the above inequality is equivalent to (4.26).

Recalling the Rodrigues formula, (4.9), we can represent P_l^m with positive m, in the form

$$P_l^m = (1-x^2)^{m/2}\frac{d^m}{dx^m}P_l(x), \qquad m \geqslant 0 \quad (4.27)$$

In potential theory m is usually taken to be positive, and (4.27) is considered the definition of P_l^m.

---

[8] If ψ is simultaneously the eigenfunction of $\mathbf{m}^2$ and m_z, $\hbar^2 l(l+1) =$
$= \int \bar{\psi}\left(m_x^2 + m_y^2 + m_z^2\right)\psi \, d\tau = \hbar^2 m_z'^2 + \int \bar{\psi}\left(m_x^2 + m_y^2\right)\psi \, d\tau \geqslant \hbar^2 m_z'^2.$

When we used Eq. (4.6), we assumed that m was an integer. In some problems connected with Pauli's and Dirac's theories of the electron, however, m can be a half-integer. The function in (4.1) then will not be a single-valued function of position, since it will change sign when φ increases by 2π. But the expressions (4.17) for Θ retain their meaning when both l and m are half-integers, which means that even in this case we can use them.

5. Some properties of spherical harmonics

Henceforth we will need to use various properties of spherical harmonics. So let us consider these properties in detail.

Recall the Cauchy formula

$$f^{(l)}(x) = \frac{l!}{2\pi i} \int \frac{f(z)}{(z-x)^{l+1}}\, dz \tag{5.1}$$

for the lth derivative of an analytic function. If we put

$$f(z) = \frac{(z^2-1)^l}{2^l l!}$$

we can represent the Legendre polynomial, defined by (4.9), in the form of an integral:

$$P_l(x) = \frac{1}{2^l}\frac{1}{2\pi i} \int \frac{(z^2-1)^l}{(z-x)^{l+1}}\, dz \tag{5.2}$$

Now we introduce a new variable ζ:

$$\frac{z-x}{z^2-1} = \frac{\zeta}{2}$$

Solving this for z and taking the root for which $z = x$ at $\zeta = 0$, we get

$$z = \frac{1}{\zeta}\left[1 - (1 - 2x\zeta + \zeta^2)^{1/2}\right]$$

It follows that

$$\frac{dz}{z-x} = \frac{d\zeta}{\zeta(1 - 2x\zeta + \zeta^2)^{1/2}}$$

which turns (5.2) into

$$P_l(x) = \frac{1}{2\pi i} \int \frac{1}{(1 - 2x\zeta + \zeta^2)^{1/2}} \frac{d\zeta}{\zeta^{l+1}} \tag{5.3}$$

If we then apply the Cauchy formula (5.1), we get

$$P_l(x) = \frac{1}{l!}\left(\frac{d^l}{d\zeta^l} \frac{1}{(1 - 2x\zeta + \zeta^2)^{1/2}}\right)_{\zeta=0} \tag{5.4}$$

Hence the $P_l(x)$ are the expansion coefficients in the Taylor series

$$\frac{1}{(1 - 2xr + r^2)^{1/2}} = \sum_{l=0}^{\infty} r^l P_l(x) \qquad (5.5)$$

which is convenient for deriving various properties of the Legendre polynomials.

For instance, let us differentiate (5.5) with respect to r. We get

$$\frac{x - r}{(1 - 2xr + r^2)^{3/2}} = \sum_{l=0}^{\infty} l r^{l-1} P_l(x) \qquad (5.6)$$

If we then multiply (5.6) by $2r$ and add the product to (5.5), we get

$$\frac{1 - r^2}{(1 - 2xr + r^2)^{3/2}} = \sum_{l=0}^{\infty} (2l + 1) r^l P_l(x) \qquad (5.7)$$

On the other hand, if we multiply (2.6) by r^2 and (2.5) by r and add the two products, we find that

$$\frac{r - xr^2}{(1 - 2xr + r^2)^{3/2}} = \sum_{l=0}^{\infty} l r^l P_{l-1}(x) \qquad (5.8)$$

But the sum of (5.6) and (5.8) is equal to (5.7) multiplied by x:

$$\sum_{l=0}^{\infty} r^l (l + 1) P_{l+1}(x) + \sum_{l=0}^{\infty} r^l l P_{l-1}(x) = \sum_{l=0}^{\infty} r^l (2l + 1) x P_l(x)$$

Identifying powers of r finally yields

$$(2l + 1) x P_l(x) = (l + 1) P_{l+1}(x) + l P_{l-1}(x) \qquad (5.9)$$

What we have arrived at is a *recursion relation* that enables us to find $P_{l+1}(x)$ if $P_l(x)$ and $P_{l-1}(x)$ are known.

Now we differentiate (5.5) with respect to x and divide the result by r. We get

$$\frac{1}{(1 - 2xr + r^2)^{3/2}} = \sum_{l=0}^{\infty} r^l \frac{dP_{l+1}}{dx}$$

If we then multiply this expression by $1 - r^2$, we will obtain

$$\frac{1 - r^2}{(1 - 2xr + r^2)^{3/2}} = \sum_{l=0}^{\infty} r^l \left(\frac{dP_{l+1}}{dx} - \frac{dP_{l-1}}{dx} \right) \qquad (5.10)$$

Comparison of (5.10) and (5.7) yields another property of the Legendre polynomials:

$$(2l + 1) P_l(x) = \frac{dP_{l+1}}{dx} - \frac{dP_{l-1}}{dx} \tag{5.11}$$

We will generalize (5.9) and (5.11) to include spherical harmonics (more correctly, the associated Legendre polynomials of the first kind). If we differentiate (5.11) m times with respect to x and then multiply the result by $(1 - x^2)^{(m+1)/2}$, with the help of (4.27) we can write the recursion relation

$$(2l + 1)\left(1 - x^2\right)^{1/2} P_l^m(x) = P_{l+1}^{m+1}(x) - P_{l-1}^{m+1}(x) \tag{5.12}$$

On the other hand, if we differentiate (5.9) m times with respect to x and multiply the result by $(1 - x^2)^{m/2}$, we get

$$(2l + 1) x P_l^m(x) + (2l + 1) m \left(1 - x^2\right)^{1/2} P_l^{m-1}(x)$$
$$= (l + 1) P_{l+1}^m(x) + l P_{l-1}^m(x)$$

The last step is to substitute $P_{l+1}^m(x) - P_{l-1}^m(x)$ for $(2l + 1) \times (1 - x^2)^{1/2} P_l^{m-1}(x)$ by using (5.12). This yields

$$(2l + 1) x P_l^m(x) = (l - m + 1) P_{l+1}^m(x) + (l + m) P_{l-1}^m(x) \tag{5.13}$$

which interrelates three spherical harmonics in succession with the same value of m.

We note that (5.12) and (5.13) hold for both positive and negative values of m. If the absolute value of the upper index of the spherical harmonic is greater than the lower index, the spherical harmonic vanishes.

Finally, let us elaborate on the system of differential equations for spherical harmonics since we will need to use it when we come to Dirac's theory of the electron.

To start with, we multiply (4.22) by $(1 - x^2)^{m/2}$ and differentiate with respect to x. We write the result as follows:

$$\frac{d}{dx}\left[\left(1 - x^2\right)^{m/2} P_l^m(x)\right]$$
$$= -(l + m)(l - m + 1)\left(1 - x^2\right)^{(m-1)/2} P_l^{m-1}(x)$$

or, if we change m to $m + 1$,

$$\frac{d}{dx}\left[(1 - x^2)^{(m+1)/2} P_l^{m+1}(x)\right]$$
$$= -(l + m + 1)(l - m)\left(1 - x^2\right)^{m/2} P_l^m(x) \tag{5.14}$$

By multiplying (4.21) into $(1 - x^2)^{-m/2}$ and differentiating with respect to x, we get

$$\frac{d}{dx}\left[\left(1 - x^2\right)^{-m/2} P_l^m(x)\right] = \left(1 - x^2\right)^{-(m+1)/2} P_l^{m+1}(x) \tag{5.15}$$

The two equations, (5.14) and (5.15), constitute a system of equations, which, after excluding $P_l^{m+1}(x)$, yields an equation for $P_l^m(x)$, (4.23). Let us replace x in (5.14) and (5.15) by a new variable: $\theta = \text{arc cos } x$. The equations then read

$$\frac{d}{d\theta}\left[(\sin\theta)^{m+1} P_l^{m+1}(\cos\theta)\right]$$

$$= (l + m + 1)(l - m)(\sin\theta)^{m+1} P_l^m(\cos\theta) \quad (5.16)$$

$$\frac{d}{d\theta}\left[(\sin\theta)^{-m} P_l^m(\cos\theta)\right] = -(\sin\theta)^{-m} P_l^{m+1}(\cos\theta) \quad (5.17)$$

or, if we differentiate explicitly,

$$\frac{d}{d\theta} P_l^m(\cos\theta) - m\cot\theta\, P_l^m(\cos\theta) = -P_l^{m+1}(\cos\theta) \quad (5.18)$$

$$\frac{d}{d\theta} P_l^{m+1}(\cos\theta) + (m+1)\cot\theta\, P_l^{m+1}(\cos\theta)$$

$$= (l + m + 1)(l - m) P_l^m(\cos\theta) \quad (5.19)$$

We will return to these equations in Dirac's theory.

6. Normalized spherical harmonics

The functions $P_l^m(x)$ of the previous sections constitute a complete set of eigenfunctions, the eigenfunctions of a hermitian operator on the left-hand side of the equation

$$-\frac{d}{dx}\left[(1 - x^2)\frac{d\Theta}{dx}\right] + \frac{m^2}{1 - x^2}\Theta = l(l+1)\Theta \quad (6.1)$$

They are orthogonal

$$\int_{-1}^{1} P_l^m(x) P_{l'}^m(x)\, dx = 0, \qquad l \neq l'$$

but not normalized. We introduce new functions

$$P_l^{*m}(x) = c_{lm} P_l^m(x) \quad (6.2)$$

and normalize them so that

$$\frac{1}{2}\int_{-1}^{+1} P_l^{*m}(x) P_{l'}^{*m}(x)\, dx = \delta_{ll'} \quad (6.3)$$

We have still to find the factors c_{lm}. But we see that

$$\frac{2}{(c_{lm})^2} = \int_{-1}^{+1} \left[P_l^m(x)\right]^2 dx \quad (6.4)$$

To evaluate the integral we substitute the product of (4.21) and (4.22) for P_l^m. We get

$$\frac{2}{(c_{lm})^2} = (-1)^m \frac{(l+m)!}{(l-m)!} \frac{1}{(2^l l!)^2} \int_{-1}^{+1} \frac{d^{l-m}(x^2-1)^l}{dx^{l-m}} \frac{d^{l+m}(x^2-1)^l}{dx^{l+m}} dx$$

If we then integrate the last expression $l - m$ times by parts and note that

$$\frac{d^{2l}(x^2-1)^l}{dx^{2l}} = (2l)!$$

we arrive at

$$\frac{2}{(c_{lm})^2} = \frac{(l+m)!}{(l-m)!} \frac{(2l)!}{(2^l l!)^2} \int_{-1}^{+1} (1-x^2)^l dx$$

This integral is easily found; it is

$$\int_{-1}^{+1} (1-x^2)^l dx = \frac{2}{2l+1} \frac{(2^l l!)^2}{(2l!)}$$

Hence

$$\frac{2}{(c_{lm})^2} = \int_{-1}^{+1} [P_l^m(x)]^2 dx = \frac{2}{2l+1} \frac{(l+m)!}{(l-m)!} \tag{6.5}$$

which implies that

$$c_{lm} = (2l+1)^{1/2} \left(\frac{(l-m)!}{(l+m)!} \right)^{1/2} \tag{6.6}$$

The normalized spherical harmonics are

$$P_l^{*m}(x) = (2l+1)^{1/2} \left(\frac{(l-m)!}{(l+m)!} \right)^{1/2} P_l^m(x) \tag{6.7}$$

Let us express these functions through the derivatives explicitly. Recalling (4.21) and (4.22), we get

$$P_l^{*m}(x) = (2l+1)^{1/2} \left(\frac{(l-m)!}{(l+m)!} \right)^{1/2} (1-x^2)^{m/2} \frac{d^{l+m}}{dx^{l+m}} \frac{(x^2-1)^l}{2^l l!} \tag{6.8}$$

and

$$P_l^{*m}(x) = (-1)^m (2l+1)^{1/2} \left(\frac{(l+m)!}{(l-m)!} \right)^{1/2} (1-x^2)^{-m/2} \frac{d^{l-m}}{dx^{l-m}} \frac{(x^2-1)^l}{2^l l!} \tag{6.9}$$

Hence we see that

$$P_l^{*-m}(x) = (-1)^m P_l^{*m}(x) \tag{6.10}$$

For the normalized spherical harmonics the recursion relations (5.12) and (5.13) read

$$xP_l^{*m}(x) = \left(\frac{l^2 - m^2}{4l^2 - 1}\right)^{1/2} P_{l-1}^{*m}(x) + \left(\frac{(l+1)^2 - m^2}{4(l+1)^2 - 1}\right)^{1/2} P_{l+1}^{*m}(x) \quad (6.11)$$

$$(1 - x^2)^{1/2} P_l^{*m}(x) = -\left(\frac{(l-m)(l-m-1)}{4l^2 - 1}\right)^{1/2} P_{l-1}^{*m+1}(x)$$

$$+ \left(\frac{(l+m+1)(l+m+2)}{4(l+1)^2 - 1}\right)^{1/2} P_{l+1}^{*m+1}(x) \quad (6.12)$$

In conclusion let us list some of the Legendre polynomials and the functions $P_l^{*m}(x)$:

$$P_0(x) = 1, \qquad P_1(x) = x, \qquad P_2(x) = \frac{1}{2}(3x^2 - 1)$$

$$P_3(x) = \frac{1}{2}(5x^3 - 3x), \qquad P_4(x) = \frac{1}{8}(35x^4 - 30x^2 + 3)$$

$$P_1^{*1}(x) = \sqrt{\frac{3}{2}}(1 - x^2)^{1/2}$$

$$P_2^{*2}(x) = \sqrt{\frac{15}{8}}(1 - x^2), \qquad P_2^{*1}(x) = \sqrt{\frac{15}{2}}(1 - x^2)^{1/2}x$$

$$P_3^{*3}(x) = \sqrt{\frac{35}{16}}(1 - x^2)^{3/2}, \qquad P_3^{*2}(x) = \sqrt{\frac{105}{8}}(1 - x^2)x$$

$$P_3^{*1}(x) = \sqrt{\frac{21}{16}}(1 - x^2)^{1/2}(5x^2 - 1)$$

7. The radial functions. A general survey

Let us consider differential equation (3.16) for the radial functions, which for the sake of convenience we will restate

$$\frac{d^2R}{dr^2} + \frac{2}{r}\frac{dR}{dr} - \frac{l(l+1)}{r^2}R + \frac{2m}{\hbar^2}[E - U(r)]R = 0 \quad (7.1)$$

To analyze this equation we must assume a certain behaviour of the potential energy for small and great distances from the nucleus. We start with the case of great distances. Assume that when $r \to \infty$ the potential energy can be represented as

$$U(r) = -\frac{A}{r} + \frac{B}{r^2} + \ldots \quad (7.2)$$

The first term on the right-hand side, $-A/r$, is the Coulomb field, which acts over great distances. Coefficient A for a valence electron is

$$A = Z^* e^2$$

with $Z^* e$ the effective charge of the nucleus (the algebraic sum of the charges of the nucleus and the inner-shell electrons), which

implies that A is positive (attraction). The wave equation for an α-particle will have a negative A (repulsion).

Let us try to find the behaviour of the solution at large values of r. For this we assume that

$$R = r^{\beta}e^{\alpha r}\left(1 + \frac{C}{r} + \dots\right) \qquad (7.3)$$

where the omitted terms are of the order of $1/r^2$ and higher. The other terms in Eq. (7.1) are

$$\frac{d^2R}{dr^2} = r^{\beta}e^{\alpha r}\left(\alpha^2 + \frac{2\alpha\beta + C\alpha^2}{r} + \dots\right)$$

$$\frac{2}{r}\frac{dR}{dr} = r^{\beta}e^{\alpha r}\left(\frac{2\alpha}{r} + \dots\right)$$

$$\frac{2m}{\hbar^2}\left(E + \frac{A}{r} + \dots\right)R = r^{\beta}e^{\alpha r}\frac{2m}{\hbar^2}\left(E + \frac{CE + A}{r} + \dots\right)$$

If we then substitute these expressions into the equation and divide $r^{\beta}e^{\alpha r}$ out, we get

$$\alpha^2 + \frac{2m}{\hbar^2}E + \left[2(\beta + 1)\alpha + \frac{2m}{\hbar^2}A + C\left(\alpha^2 + \frac{2m}{\hbar^2}E\right)\right]\frac{1}{r} + \dots = 0$$

from which we derive two equations

$$\alpha^2 + \frac{2m}{\hbar^2}E = 0$$

$$2(\beta + 1)\alpha + \frac{2m}{\hbar^2}A = 0 \qquad (7.4)$$

for the constants α and β. They have two solutions

$$\alpha = (-2mE/\hbar^2)^{1/2} \qquad (7.5)$$

$$\beta = -1 + A\alpha/(2E) \qquad (7.6)$$

corresponding to the two possible signs in (7.5). If we choose one sign, the leading terms of the solution to (7.1) are

$$R = \frac{1}{r}\left\{C_1\exp\left[\alpha\left(r + \frac{A}{2E}\log r\right)\right] + C_2\exp\left[-\alpha\left(r + \frac{A}{2E}\log r\right)\right]\right\} \qquad (7.7)$$

We see that the behaviour of the solution is different depending on whether α is real or imaginary.

If the total energy, E, is positive (which in classical mechanics corresponds to orbits that extend to infinity), α is pure imaginary:

$$E > 0, \qquad \alpha = i(2mE/\hbar^2)^{1/2} \qquad (7.8)$$

Then the general solution of (7.7) will, as $r \to \infty$, alternate signs and tend to zero like $1/r$. However, it will fall off so slowly that

the integral

$$\int_{r_0}^{\infty} r^2 |R(r)|^2 \, dr \qquad (7.9)$$

with r_0 a certain finite constant, will diverge.

If the total energy is negative, α is real (we will assume it to be positive):

$$E < 0, \qquad \alpha = |(-2mE/\hbar^2)^{1/2}| \qquad (7.10)$$

and the behaviour of the solution depends on whether C_1 vanishes. If $C_1 \neq 0$, then (7.7) will, as $r \to \infty$, grow infinitely. But if C_1 is zero, R will fall off at infinity according to the exponential law and (7.9) will converge.

What remains is to study the case when $E = 0$. We must look for the solution in a somewhat different form, namely

$$R = r^{\beta_1} e^{\alpha_1 r^{1/2}} \left(1 + \frac{C}{r^{1/2}} + \ldots \right) \qquad (7.11)$$

Proceeding as we did above, we get

$$\frac{\alpha_1^2}{4} + \frac{2mA}{\hbar^2} + \left[\alpha_1 \left(\beta_1 + \frac{3}{4} \right) + C \left(\frac{\alpha_1^2}{4} + \frac{2mA}{\hbar^2} \right) \right] \frac{1}{r^{1/2}} + \ldots = 0$$

which implies that

$$\alpha_1 = 2 \, (-2mA/\hbar^2)^{1/2} \qquad (7.12)$$
$$\beta_1 = -\, {}^3/_4 \qquad (7.13)$$

Hence the general solution in this case is

$$R = r^{-3/4} \left(C_1' e^{\alpha_1 r^{1/2}} + C_2' e^{-\alpha_1 r^{1/2}} \right) \qquad (7.14)$$

If $A > 0$ (attraction at great distances), α_1 is pure imaginary and R will be finite. If $A < 0$ (repulsion), R will, generally speaking, grow as $r \to \infty$.

We now turn to the case of r small [in Eq. (7.1)]. Let us assume that at $r = 0$ the potential energy grows no faster than $1/r$:

$$U(r) = -\frac{A_1}{r} + (\text{finite function for } r = 0) \qquad (7.15)$$

This corresponds to a Coulomb field at small distances from the nucleus. The coefficient A_1 can differ from A in (7.2). For a valence electron

$$A_1 = Ze^2$$

with Ze the charge of the nucleus.

12—2186

We seek the solution in the form

$$R = r^\alpha + Cr^{\alpha+1} + \ldots \qquad (7.16)$$

Substituting (7.16) into (7.1) and equating with zero the term that has r in the lowest power, we get

$$\alpha(\alpha+1) - l(l+1) = 0$$

which simply means that

$$\alpha = l \quad \text{or} \quad \alpha = -l-1 \qquad (7.17)$$

The general solution of our equation will have the following form:

$$R = C'r^l(1 + a'r + \ldots) + C''r^{-l-1}(1 + a''r + \ldots) \quad (7.18)$$

Thus at $r = 0$ the solution does not depend either on E or on the expansion coefficients of $U(r)$ in (7.15), and we must put $C'' = 0$ if we want our solution to remain finite at $r = 0$.

Combining this result with the one obtained for large values of r, we come to the following conclusion.

At $E > 0$ all solutions, including the one that remains finite at $r = 0$, vanish as $r \to \infty$. Then to find a solution $R(r)$ that is finite everywhere it is sufficient to take the solution that at $r = 0$ is finite, which means that *at $E > 0$ the Hamiltonian has a continuous spectrum in the interval from 0 to ∞* (the eigenvalue $E = 0$ lies in the continuous spectrum only for attractive forces). At the same time for $E \geqslant 0$ the integral (7.9) is not finite. This implies that at $E \geqslant 0$ *there cannot be a discrete spectrum*, since eigenfunctions belonging to a discrete spectrum are square integrable.

Now what happens when $E < 0$? The solution that at $r = 0$ remains finite passes, as $r \to \infty$, into an expression of type (7.7) with a real exponent α. The ratio C_1/C_2 is determined by the value of E. Two cases are possible here. In one this ratio is finite and $R(r)$ tends to infinity as $r \to \infty$, which means that the corresponding value of E is not an eigenvalue of the Hamiltonian. In the other this ratio is zero and $R(r)$ falls off at infinity so rapidly that integral (7.9) is finite. The corresponding $E = E_n$ is an eigenvalue belonging to the discrete spectrum, and we conclude that *at $E < 0$ there is no continuous spectrum, and either there is a discrete spectrum or there is no spectrum at all.* The first case holds for attraction, and the second for repulsion.

Thus for attractive forces the eigenvalue spectrum of the Hamiltonian consists of a collection of negative numbers

$$E_1, E_2, \ldots, E_n, \ldots \text{ (the discrete spectrum)} \qquad (7.19)$$

and a continuous set of values

$$0 \leqslant E < \infty \text{ (the continuous spectrum)} \qquad (7.20)$$

Since the radial equation contains l as a parameter, the eigenvalues belonging to the discrete spectrum also depend on l, and we will denote them by E_{nl}. For the corresponding radial functions we will use R_{nl} for the discrete spectrum and R_{El} for the continuous spectrum.

Let us consider an electron that is in a state with definite energy and angular momentum. According to the probabilistic interpretation of the wave function, the relative probability of the electron's being (after the proper measuring process has been completed) at a distance between r and $r + dr$ is $|R_{nl}(r)|^2 r^2\, dr$ for the discrete spectrum ($E_{nl} < 0$) and $|R_{El}(r)|^2 r^2\, dr$ for the continuous spectrum ($E > 0$). In classical mechanics, on the other hand, negative energies correspond to closed orbits and positive energies to orbits that extend to infinity. Using the classical-analogy concept, we can expect that for the discrete spectrum the probability of finding the electron at great distances from the atom is considerably smaller than for the continuous spectrum. Indeed, if we recall the properties of the radial functions, we see that $|R_{nl}|^2 r^2$ falls off at infinity according to the exponential law, whereas $|R_{El}|^2 r^2$ usually remains finite.

8. Description of the states of a valence electron. Quantum numbers

Let us summarize what we have learned about an electron moving in a field with central symmetry (a valence electron in an atom). Such electrons can be described by a wave function

$$\psi_{nlm} = e^{-iE_{nl}t/\hbar} R_{nl}(r) Y_{lm}(\theta, \varphi) \qquad (8.1)$$

for the discrete spectrum, and by a wave function

$$\psi_{Elm} = e^{-iEt/\hbar} R_{El}(r) Y_{lm}(\theta, \varphi) \qquad (8.2)$$

for the continuous spectrum. We assume the radial functions to be normalized. Namely, for the discrete spectrum

$$\int_0^\infty |R_{nl}(r)|^2 r^2\, dr = 1 \qquad (8.3)$$

and for the continuous spectrum

$$\lim_{\Delta E \to 0} \frac{1}{\Delta E} \int_0^\infty \left| \int_E^{E+\Delta E} R_{El}(r)\, dE \right|^2 r^2\, dr = 1 \qquad (8.4)$$

The spherical harmonic $Y_{lm}(\theta, \varphi)$ is expressed, according to the results of Sections 4 and 6, in the following way:

$$Y_{lm}(\theta, \varphi) = \frac{1}{(4\pi)^{1/2}} e^{im\varphi} P_l^{*m}(\cos\theta) \qquad (8.5)$$

and the normalization condition is

$$\int_{\theta=0}^{\pi} \int_{\varphi=0}^{2\pi} |Y_{lm}(\theta, \varphi)|^2 \sin\theta \, d\theta \, d\varphi = 1 \qquad (8.6)$$

Functions (8.1) and (8.2) are the simultaneous eigenfunctions of the Hamiltonian H, the square of angular momentum $\mathbf{m}^2$, and the component of angular momentum m_z along the z axis. This is why in a state described by either (8.1) or (8.2) these three quantities have definite values simultaneously:

quantity	eigenvalue
H	E_{nl} or E
$\mathbf{m}^2$	$l(l+1)\,\hbar^2$
m_z	$m\hbar$

$$\qquad\qquad\qquad (8.7)$$

Thus a state is characterized by three *quantum numbers, n, l,* and *m*, or by one continuous parameter, E, and two quantum numbers, l and m. The label n is called the *principal quantum number*; it is usually defined as the sum

$$n = n_r + l + 1 \qquad (8.8)$$

where n_r determines the number of *zeros* (or *nodes*) of $R_{nl}(r)$. The label n_r is called the *radial quantum number*, and l the *azimuthal quantum number.* Such a definition for n is possible because when an eigenfunction of a differential operator of type in (7.1) belongs to a discrete spectrum, it is characterized by the numbers of its nodes. Since n_r is nonnegative, the principal quantum number is greater than the azimuthal quantum number at least by one.

If we study Eq. (7.1), we see that m is not included. This means that energy levels E_{nl} do not depend on m (in other words, the *energy state term* does not determine the value of m). This was to be expected since $m\hbar$ is the value of the angular momentum component along the z axis, and in the case of a central field the direction of this (and any other) axis is undefined from the physical point of view. But if a magnetic field is applied to the system [9] (we call the direction of this field the *z axis*), the energy levels

---

[9] As we noted at the beginning of this chapter, we must use the Dirac equation in the case of a magnetic field.

will depend on m as well. For this reason m is called the *magnetic quantum number.*

The Coulomb field has a specific property, which we will study in the next chapter — the energy of an electron in such a field depends only on n.

For the general case of a central field each energy level E_{nl} has corresponding to it $2l + 1$ eigenfunctions, which can be found if in (8.1) we put

$$m = -l, \ -l+1, \ \ldots, \ l-1, \ l \tag{8.9}$$

Because of this the energy level is $(2l + 1)$-fold degenerate. In the case of a Coulomb field the degeneracy is greater, since for a given n the azimuthal quantum number, l, can take the values

$$l = 0, \ 1, \ \ldots, \ n-1 \tag{8.10}$$

and to each value of l there correspond $2l + 1$ different values of m. The total degree of degeneracy is the sum

$$1 + 3 + 5 + \ldots + 2n - 1 = n^2 \tag{8.11}$$

There is an accepted spectroscopic notation for energy state terms. Terms with the same value of n but different values of l are denoted by Latin letters $s, p, d, f, g, \ldots$. For instance

term	spectroscopic notation
$n = 1, \ l = 0$	(1s)
$n = 2, \ l = 0$	(2s)
$n = 2, \ l = 1$	(2p)
$n = 3, \ l = 0$	(3s)
$n = 3, \ l = 1$	(3p)
$n = 3, \ l = 2$	(3d)

We note that according to Schrödinger's theory no spectral term is degenerate, whereas in reality all terms except the s terms (corresponding to $l = 0$) are *doublets*, that is, consist of two adjacent terms (the *fine structure of spectral lines*). As we will see in Part V, Dirac's theory of the electron provides an explanation of this phenomenon.

9. The selection rule

Not knowing the exact form of the radial functions, we cannot calculate the elements of Heisenberg's matrices, which according to the results obtained in Section 3, Chapter III, characterize the intensity of the spectral lines corresponding to various transitions.

However, we can state which of these matrix elements vanish, because we know the angular dependence of the eigenfunctions. In other words, we can evolve the selection rule.

First we must generalize the formulas for the intensities derived in Sections 3 and 4, Chapter III, to the case of several quantum numbers and degenerate energy levels. We recall that

$$I_{nn'} = e^2 \omega_{nn'}^4 \left(|x_{nn'}|^2 + |y_{nn'}|^2 + |z_{nn'}|^2 \right) \tag{9.1}$$

for the discrete spectrum, and

$$I_n(E) \, \Delta E = e^2 \omega_n^4(E) \left[|(E_n|x|E)|^2 + |(E_n|y|E)|^2 \right.$$
$$\left. + |(E_n|z|E)|^2 \right] \Delta E \tag{9.2}$$

for the continuous spectrum. In our case the state of an electron is given by specifying three quantum numbers. Hence the elements of Heisenberg's matrix for x will be

$$x_{nn'} = (nlm|x|n'l'm') \tag{9.3}$$

for the discrete spectrum, and

$$(E_n|x|E) = (nlm|x|El'm') \tag{9.4}$$

for the continuous spectrum. The frequencies $\omega_{nn'}$ and $\omega_n(E)$ are understood to be respectively

$$\omega_{nn'} = (E_{nl} - E_{n'l'})/\hbar \tag{9.5}$$

$$\omega_n(E) = (E_{nl} - E)/\hbar \tag{9.6}$$

or, more correctly, the absolute values of these quantities.

Transitions that differ in the values of the quantum numbers m and m' may have the same frequencies. In normal conditions (without a magnetic field) such transitions cannot be distinguished. What is observed is the sum of the intensities of all the transitions that have the same frequency. Thus we must change $|x_{nn'}|^2$ in (9.1):

$$|x_{nn'}|^2 \rightarrow \sum_{m=-l}^{l} \sum_{m'=-l'}^{l'} |(nlm|x|n'l'm')|^2 \tag{9.7}$$

The same holds for y and z. In the continuous spectrum the values of the energy parameter form a continuum, which means that one cannot specify the quantum numbers l' and m' by simply stating the value of E. Hence we must change $|(E_n|x|E)|^2$ in (9.2):

$$|(E_n|x|E)|^2 \rightarrow \sum_{m=-l}^{l} \sum_{l'=0}^{\infty} \sum_{m=-l'}^{l'} |(nlm|x|El'm')|^2 \tag{9.8}$$

We note that because of the selection rule, which we derive below, the sum (9.8) contains only a finite number of terms.

With the given changes formulas (9.1) and (9.2) hold for the considered general case.

If magnetic field is present, one can distinguish the transitions with different values of m. In this case the separate terms in (9.7) may be of interest.

We now turn to the actual evaluation of the elements of Heisenberg's matrices corresponding to the coordinates x, y, z for the discrete spectrum, namely

$$(nlm \mid x \mid n'l'm') = \iiint r \sin\theta \cos\varphi \, \bar{\psi}_{nlm} \psi_{n'l'm'} r^2 \sin\theta \, d\theta \, d\varphi \, dr \qquad (9.9)$$

$$(nlm \mid y \mid n'l'm') = \iiint r \sin\theta \sin\varphi \, \bar{\psi}_{nlm} \psi_{n'l'm'} r^2 \sin\theta \, d\theta \, d\varphi \, dr \qquad (9.10)$$

$$(nlm \mid z \mid n'l'm') = \iiint r \cos\theta \, \bar{\psi}_{nlm} \psi_{n'l'm'} r^2 \sin\theta \, d\theta \, d\varphi \, dr \qquad (9.11)$$

where because of (8.1).

$$\bar{\psi}_{nlm} \psi_{n'l'm'} = e^{i\omega t} \bar{R}_{nl} R_{n'l'} \bar{Y}_{lm} Y_{l'm'} \qquad (9.12)$$

or

$$\bar{\psi}_{nlm} \psi_{n'l'm'} = e^{i\omega t} \bar{R}_{nl} R_{n'l'} \frac{1}{4\pi} P_l^{*m} P_{l'}^{*m'} e^{i(m'-m)\varphi} \qquad (9.13)$$

with ω being actually as in (9.5).

Each of the triple integrals factors into three simple integrals, and the integrals with respect to r in (9.9), (9.10), and (9.11) are the same, namely

$$r(nl; \; n'l') = \int_0^\infty \bar{R}_{nl} R_{n'l'} r^3 \, dr \qquad (9.14)$$

We denote the integrals with respect to θ and φ as follows:

$$(lm \mid \sin\theta \cos\varphi \mid l'm') = \frac{1}{4\pi} \iint P_l^{*m} P_{l'}^{*m'} e^{i(m'-m)\varphi} \sin^2\theta \cos\varphi \, d\theta \, d\varphi \qquad (9.15)$$

$$(lm \mid \sin\theta \sin\varphi \mid l'm') = \frac{1}{4\pi} \iint P_l^{*m} P_{l'}^{*m'} e^{i(m'-m)\varphi} \sin^2\theta \sin\varphi \, d\theta \, d\varphi \qquad (9.16)$$

$$(lm \mid \cos\theta \mid l'm') = \frac{1}{4\pi} \iint P_l^{*m} P_{l'}^{*m'} e^{i(m'-m)\varphi} \sin\theta \cos\theta \, d\theta \, d\varphi \qquad (9.17)$$

Thus the matrix elements for x, y, z will be equal (apart from a constant factor $e^{i\omega t}$) to the products of (9.14) by (9.15), (9.16), and (9.17) respectively. For a continuous spectrum all the formulas will remain the same with one obvious exception: we must change $R_{n'l'}$ to $R_{El'}$. Then

$$r(nl; \; El') = \int_0^\infty \bar{R}_{nl} R_{El'} r^3 \, dr \qquad (9.14^*)$$

Let us evaluate the integrals in (9.15), (9.16), and (9.17). Since these integrals are factors in (9.9), (9.10), and (9.11), we see that if for some values of l, m, l', m' they vanish, the corresponding elements of Heisenberg's matrices vanish too. This is the selection rule.

We start with integral (9.17) as the simpler one. Integrating with respect to φ, we see that it is nonzero only if $m = m'$ since

$$\frac{1}{2\pi} \int_0^{2\pi} e^{i(m'-m)\varphi} \, d\varphi = \delta_{mm'} \tag{9.18}$$

To evaluate the integral with respect to θ we introduce a new variable

$$x = \cos\theta$$

Then (9.17) reads

$$(lm \mid \cos\theta \mid l'm') = \delta_{mm'} \frac{1}{2} \int_{-1}^{1} P_l^{*m}(x) P_{l'}^{*m}(x) x \, dx \tag{9.19}$$

If we use the recursion relation (6.11), we can change $xP_l^{*m}(x)$ to

$$xP_l^{*m}(x) = \left(\frac{(l+1)^2 - m^2}{4(l+1)^2 - 1} \right)^{1/2} P_{l+1}^{*m}(x) + \left(\frac{l^2 - m^2}{4l^2 - 1} \right)^{1/2} P_{l-1}^{*m}(x) \tag{9.20}$$

Bringing in the orthogonality and normalization conditions for $P_l^{*m}(x)$ yields

$$(lm \mid \cos\theta \mid l'm') = \delta_{mm,} \left[\left(\frac{(l+1)^2 - m^2}{4(l+1)^2 - 1} \right)^{1/2} \delta_{l+1\,l'} + \right.$$
$$\left. + \left(\frac{l^2 - m^2}{4l^2 - 1} \right)^{1/2} \delta_{l\,l'+1} \right] \tag{9.21}$$

Hence a matrix element does not vanish only if $l - l' = \pm 1$.

To evaluate the integrals (9.15) and (9.16) it is convenient to build their linear combination

$$(lm \mid \sin\theta\, e^{i\varphi} \mid l'm') = (lm \mid \sin\theta \cos\varphi \mid l'm')$$
$$+ i\,(lm \mid \sin\theta \sin\varphi \mid l'm') \tag{9.22}$$

which we can use to find the expressions for (9.15) and (9.16) in the following way:

$$(lm \mid \sin\theta \cos\varphi \mid l'm') = \frac{1}{2} [(lm \mid \sin\theta\, e^{i\varphi} \mid l'm') + \overline{(l'm' \mid \sin\theta\, e^{i\varphi} \mid lm)}] \tag{9.23}$$

$$(lm \mid \sin\theta \sin\varphi \mid l'm') = \frac{1}{2i} [(lm \mid \sin\theta\, e^{i\varphi} \mid l'm')$$
$$- \overline{(l'm' \mid \sin\theta\, e^{i\varphi} \mid lm)}] \tag{9.24}$$

In explicit form, (9.22) is

$$(lm \mid \sin \theta \, e^{i\varphi} \mid l'm') = \frac{1}{4\pi} \iint P_l^{*m} P_{l'}^{*m'} e^{i \, (m'-m+1) \, \varphi} \sin^2 \theta \, d\theta \, d\varphi \qquad (9.25)$$

If we integrate with respect to φ and introduce the variable $x = \cos \theta$, we get

$$(lm \mid \sin \theta \, e^{i\varphi} \mid l'm') = \delta_{m \, m'+1} \frac{1}{2} \int_{-1}^{1} (1 - x^2)^{1/2} P_l^{*m} (x) P_{l'}^{*m-1} (x) \, dx \qquad (9.26)$$

By changing m to $m-1$ in (6.12) we find that

$$(1 - x^2)^{1/2} P_{l'}^{*m-1} (x) = \left(\frac{(l' + m) (l' + m + 1)}{4 (l' + 1)^2 - 1} \right)^{1/2} P_{l'+1}^{*m} (x)$$
$$- \left(\frac{(l' - m + 1) (l' - m)}{4 l'^2 - 1} \right)^{1/2} P_{l'-1}^{*m} (x) \qquad (9.27)$$

which we then substitute into (9.26). Applying the normalization and orthogonality conditions for $P_l^{*m} (x)$ and also expressions of type

$$f (l') \delta_{l \, l'-1} = f (l + 1) \delta_{l \, l'-1}$$

we get

$$(lm \mid \sin \theta \, e^{i\varphi} \mid l'm') = \delta_{m-1 \, m'} \left[\left(\frac{(l + m - 1) (l + m)}{4 l^2 - 1} \right)^{1/2} \delta_{l-1 \, l'} \right.$$
$$\left. - \left(\frac{(l - m + 1) (l - m + 2)}{4 (l + 1)^2 - 1} \right)^{1/2} \delta_{l+1 \, l'} \right] \qquad (9.28)$$

whence

$$\overline{(l'm' \mid \sin \theta \, e^{i\varphi} \mid lm)} = \delta_{m+1 \, m'} \left[\left(\frac{(l + m + 1) (l + m + 2)}{4 (l + 1)^2 - 1} \right)^{1/2} \delta_{l+1 \, l'} \right.$$
$$\left. - \left(\frac{(l - m) (l - m - 1)}{4 l^2 - 1} \right)^{1/2} \delta_{l-1 \, l'} \right] \qquad (9.29)$$

What remains to be calculated is the half-sum of (9.28) and (9.29) and the half-difference divided by i. According to (9.23) and (9.24) this gives us the matrix elements (9.15) and (9.16).

We see that (9.15), (9.16), and (9.17) and hence the matrix elements for x, y, z are nonzero only if

$$l - l' = \pm 1 \qquad (9.30)$$

which is the selection rule for the azimuthal quantum number, l. The rule states that transitions between s and p spectral terms, p and d terms, etc. are possible, whereas transitions between, say, s and d terms are forbidden. This agrees with experimental data.

As to the magnetic quantum number m, the elements of the matrices for coordinate z do not vanish only if

$$m - m' = 0 \qquad (9.31)$$

and the matrix elements for x and y do not vanish if

$$m - m' = \pm 1 \qquad (9.32)$$

These two conditions are the selection rule form. Transitions that satisfy (9.31) produce light that is polarized along the z axis, whereas those that satisfy (9.32) produce light that is polarized in the xy plane. As we have already mentioned, transitions that correspond to definite values of m and m' can be observed only in the presence of a magnetic field (directed along the z axis).

We present the matrix elements (9.15), (9.16), and (9.17), which correspond to various transitions, in table form:

$$l' = l - 1$$

$$(lm \mid \sin\theta \cos\varphi \mid l-1 \; m-1) = \frac{1}{2}\left(\frac{(l+m-1)(l+m)}{4l^2-1}\right)^{1/2}$$

$$(lm \mid \sin\theta \cos\varphi \mid l-1 \; m+1) = -\frac{1}{2}\left(\frac{(l-m)(l-m-1)}{4l^2-1}\right)^{1/2}$$

$$(lm \mid \sin\theta \sin\varphi \mid l-1 \; m-1) = -\frac{i}{2}\left(\frac{(l+m-1)(l+m)}{4l^2-1}\right)^{1/2}$$

$$(lm \mid \sin\theta \sin\varphi \mid l-1 \; m+1) = -\frac{i}{2}\left(\frac{(l-m)(l-m-1)}{4l^2-1}\right)^{1/2}$$

$$(lm \mid \cos\theta \mid l-1 \; m) = \left(\frac{l^2-m^2}{4l^2-1}\right)^{1/2}$$

$$l' = l + 1$$

$$(lm \mid \sin\theta \cos\varphi \mid l+1 \; m-1) = -\frac{1}{2}\left(\frac{(l-m+1)(l-m+2)}{4(l+1)^2-1}\right)^{1/2}$$

$$(lm \mid \sin\theta \cos\varphi \mid l+1 \; m+1) = \frac{1}{2}\left(\frac{(l+m+1)(l+m+2)}{4(l+1)^2-1}\right)^{1/2}$$

$$(lm \mid \sin\theta \sin\varphi \mid l+1 \; m-1) = \frac{i}{2}\left(\frac{(l-m+1)(l-m+2)}{4(l+1)^2-1}\right)^{1/2}$$

$$(lm \mid \sin\theta \sin\varphi \mid l+1 \; m+1) = \frac{i}{2}\left(\frac{(l+m+1)(l+m+2)}{4(l+1)^2-1}\right)^{1/2}$$

$$(lm \mid \cos\theta \mid l+1 \; m) = \left(\frac{(l+1)^2-m^2}{4(l+1)^2-1}\right)^{1/2}$$

Let us now build sums of type (9.7), which we will need for the intensities. With the help of the relationship

$$\sum_{m=-l}^{l} m^2 = \frac{1}{3} l(l+1)(2l+1) \tag{9.33}$$

we easily find that

$$\sum_{m, m'} |(lm|\cos\theta|l-1\,m')|^2 = \frac{1}{3}l \tag{9.34}$$

and by analogy

$$\sum_{m, m'} |(lm|\sin\theta\cos\varphi|l-1\,m')|^2 = \frac{1}{3}l \tag{9.35}$$

and

$$\sum_{m, m'} |(lm|\sin\theta\sin\varphi|l-1\,m')|^2 = \frac{1}{3}l \tag{9.36}$$

All three sums have the same value $l/3$, which was to be expected since after we exclude m (in the summing process) any specific direction in the system becomes meaningless and all three axes play the same role. The sums with $l' = l + 1$ can be found if we change l to $l + 1$, which gives $(l + 1)/3$.

Our results make it possible to find the general formula for intensities. In the discrete spectrum the transition intensities are

$$I(nl; n'\,l-1) = e^2 \left(\frac{E_{nl} - E_{n'\,l-1}}{\hbar}\right)^4 |r(nl; n'\,l-1)|^2\,l \tag{9.37}$$

$$I(nl; n'\,l+1) = e^2 \left(\frac{E_{nl} - E_{n'\,l+1}}{\hbar}\right)^4 |r(nl; n'\,l+1)|^2 (l+1) \tag{9.37*}$$

To find the intensities of transition from the continuous spectrum to the discrete one, we must take the sum of the corresponding expressions and multiply it by ΔE:

$$I(nl; E)\,\Delta E = e^2 \left(\frac{E_{nl} - E}{\hbar}\right)^4 [\,|r(nl; E\,l-1)|^2\,l$$
$$+ |r(nl; E\,l+1)|^2 (l+1)]\,\Delta E \tag{9.38}$$

Finally, in the case of a Coulomb field we must sum with respect to l since in this case E_{nl} does not depend on l.

THE COULOMB FIELD

1. General remarks

Here we will examine a particular case of the general problem considered in the previous chapter — the state of a particle that is attracted to or repelled from a fixed centre (the nucleus of an atom) via the Coulomb law. This is an interesting case because, on the one hand, certain important physical problems (the hydrogen atom, for instance) reduce to it, and, on the other, it permits an exact solution. The results that we obtained in Chapter IV, naturally, can be applied in full measure to the case of a Coulomb field. For instance, the separation of variables, the angular dependence of the wave function (spherical harmonics), and the selection rule can all be used. More than that, in the case of a Coulomb field it is possible to solve the radial equation exactly and thus find the energy levels and the intensities and frequencies of the spectral lines, thereby completing the solution.

This solution is simple enough to use as the starting approximation when a constant electric field disturbs a hydrogen atom (the Stark effect). We will consider this problem too.

Finally, the theory of the motion of a particle repelled via the Coulomb law makes it possible to deduce the Rutherford law for the scattering of α-particles from nuclei. This theory is also an interesting illustration of the probabilistic interpretation of quantum mechanics.

2. The radial equation for the hydrogen atom. Atomic units

The potential energy of the electron in the hydrogen atom, the electron being attracted via the Coulomb law to the nucleus (proton), is

$$U(r) = -\frac{e^2}{r} \tag{2.1}$$

with r the distance from the electron to the nucleus, which can be considered the origin of coordinates and fixed, since it is more massive than the electron. (The ratio of nuclear mass to electron mass for hydrogen is approximately 1836.) By Eq. (3.16), Chapter IV, we arrive at the equation for the radial functions of the

hydrogen atom:

$$\frac{d^2R}{dr^2} + \frac{2}{r}\frac{dR}{dr} - \frac{l(l+1)}{r^2}R + \frac{2m}{\hbar^2}\left(E + \frac{e^2}{r}\right)R = 0 \qquad (2.2)$$

If we were to account for the motion of the nucleus, we would have an equation of the same type, but instead of the electron mass, m, there would be the *reduced mass*

$$m' = \frac{mM}{m+M} \qquad (2.3)$$

where M is the mass of the nucleus.

Let us introduce a system of units with Planck's constant h divided by 2π, the electron charge, and the electron mass as base units:

$$\hbar = \frac{h}{2\pi} = \frac{1}{2\pi} \times 6.626 \times 10^{-27} \text{ erg s}$$

$$e = 4.80 \times 10^{-10} \text{ esu}$$

$$m = 9.11 \times 10^{-28} \text{ g} \qquad (2.4)$$

In practical terms this means that in such a "natural" system of units the unit of length is

$$a = \frac{\hbar^2}{me^2} = 0.529 \times 10^{-8} \text{ cm} \qquad (2.5)$$

and the unit of energy is

$$E_0 = \frac{me^4}{\hbar^2} = \frac{e^2}{a} = e\frac{e}{a} = 27.21 \text{ eV} \qquad (2.6)$$

called a *hartree*. The unit of velocity will then be $e^2/\hbar$, which is the 137th part of the speed of light.

With this in mind, in Eq. (2.2) we put

$$r_1 = \frac{r}{a}, \qquad \varepsilon = \frac{E}{E_0} \qquad (2.7)$$

(do not confuse with the smallness parameter in perturbation theory, ε). After this the equation reads

$$\frac{d^2R}{dr_1^2} + \frac{2}{r_1}\frac{dR}{dr_1} + \left(2\varepsilon + \frac{2}{r_1} - \frac{l(l+1)}{r_1^2}\right)R = 0 \qquad (2.8)$$

If we then substitute

$$R = \frac{1}{r_1^{1/2}}\, y \qquad (2.9)$$

into Eq. (2.8), we get

$$\frac{d^2y}{dr_1^2} + \frac{1}{r_1}\frac{dy}{dr_1} + \left(2\varepsilon + \frac{2}{r_1} - \frac{s^2}{4r_1^2}\right)y = 0 \qquad (2.10)$$

where we put

$$s = 2l + 1 \tag{2.11}$$

This equation occurs in a number of problems (to name some, Dirac's theory of the hydrogen atom, the Stark effect, the scattering of α-particles), and parameter s may not necessarily be an odd integer. This is why we will consider Eq. (2.10) in greater detail and will assume s to be not an integer but simply a non-negative quantity, $s \geqslant 0$ (obviously this can be done since s enters into the equation only through s^2).

3. Solution of an auxiliary problem

From the general analysis of the radial equation (Section 7, Chapter IV) we know that negative values of ε correspond to a discrete spectrum and positive values to a continuous spectrum. To study the discrete spectrum we introduce

$$x = r_1 (- 8\varepsilon)^{1/2} \tag{3.1}$$

as the independent variable and put

$$\lambda = (- 2\varepsilon)^{-1/2} \tag{3.2}$$

Here λ will obviously be real (we will consider it positive); x will be real and will change within the same limits as r, namely, between 0 and ∞. Equation (2.10) then reads

$$x \frac{d^2y}{dx^2} + \frac{dy}{dx} + \left(- \frac{x}{4} + \lambda - \frac{s^2}{4x} \right) y = 0 \tag{3.3}$$

or

$$- \frac{d}{dx} \left(x \frac{dy}{dx} \right) + \left(\frac{x}{4} + \frac{s^2}{4x} \right) y = \lambda y \tag{3.3*}$$

The operator on the left-hand side of (3.3*) is hermitian, and λ plays the role of a parameter. By introducing the quantities (3.1) and (3.2) we eliminate, so to say, the continuous spectrum and reduce the solution of Eq. (2.10) to the solution of an auxiliary problem, namely, the eigenvalue equation (3.3*).

Let us study the behaviour of Eq. (3.3) for very small and very large values of x. We could use the results of Section 7, Chapter IV, but it is simpler to repeat our reasoning for Eq. (3.3).

For small values of x we put

$$y = x^\alpha + ax^{\alpha+1} + \dots \tag{3.4}$$

and get for α two values:

$$\alpha = \pm s/2 \tag{3.5}$$

For large values of x we put

$$y = e^{-\alpha x} x^\beta \left(1 + \frac{a'}{x} + \ldots\right) \tag{3.6}$$

and get for both α and β two values:

$$\alpha = \frac{1}{2}, \qquad \beta = -\frac{1}{2} + \lambda \tag{3.7}$$

and

$$\alpha = -\frac{1}{2}, \qquad \beta = -\frac{1}{2} - \lambda \tag{3.7*}$$

From this we conclude that the sought solution for small values of x must be of the form

$$y = C x^{s/2} (1 + ax + \ldots) \tag{3.8}$$

and for large values of x of the form

$$y = C' e^{-x/2} x^{\lambda - 1/2} \left(1 + \frac{a'}{x} + \ldots\right) \tag{3.9}$$

Hence if we put

$$y = e^{-x/2} x^{s/2} Q(x) \tag{3.10}$$

then function $Q(x)$ must satisfy the following conditions:

$Q(x)$ is finite at $x = 0$

$Q(x)$ is of the order of $x^{\lambda - (s+1)/2}$ as $x \to \infty$ $\qquad$ (3.11)

and the equation for y, (3.3), leads us to the following equation for $Q(x)$:

$$x \frac{d^2 Q}{dx^2} + (s + 1 - x) \frac{dQ}{dx} + \left(\lambda - \frac{s+1}{2}\right) Q = 0 \tag{3.12}$$

This equation can be solved in two ways: by using power series or by using definite integrals. We will apply the first method here and use the second method when we solve an analogous equation for the continuous spectrum.

Let us seek the solution to Eq. (3.12) in the form

$$Q = \sum_{n=0}^{\infty} a_n x^n \tag{3.13}$$

When the series is substituted into the equation and the coefficients of the powers of x are equated with zero, we get the relationship

$$n(n + s) a_n + \left(-n + \lambda + \frac{1}{2} - \frac{s}{2}\right) a_{n-1} = 0 \tag{3.14}$$

which serves as an equation for successive determination of the expansion coefficients.

Coefficient a_0 remains arbitrary, and the other coefficients are expressed in terms of it in the following manner:

$$a_1 = \frac{(s+1)/2 - \lambda}{1(s+1)} a_0$$

$$a_2 = \frac{(s+1)/2 - \lambda + 1}{2(s+2)} a_1 = \frac{[(s+1)/2 - \lambda][(s+1)/2 - \lambda + 1]}{1 \times 2(s+1)(s+2)} a_0$$

$$\cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \quad (3.15)$$

Hence, if we introduce the *confluent hypergeometric function*

$$F(\alpha;\ \gamma;\ x) = 1 + \frac{\alpha}{\gamma} \frac{x}{1!} + \frac{\alpha(\alpha+1)}{\gamma(\gamma+1)} \frac{x^2}{2!} + \cdots \qquad (3.16)$$

we can write

$$Q = a_0 F\left(\frac{s+1}{2} - \lambda;\ s+1;\ x\right) \qquad (3.17)$$

Two cases are possible. If

$$\lambda = \frac{s+1}{2} + p, \qquad p = 0,\ 1,\ 2,\ \ldots \qquad (3.18)$$

then a_{p+1} and all subsequent coefficients vanish and the series in (3.13) is terminated. For Q we then have not an infinite series but a polynomial. If condition (3.18) does not hold, the series in (3.13) is an infinite one. It converges, since the ratio of any two subsequent terms,

$$\frac{a_n x^n}{a_{n-1} x^{n-1}} = x \frac{n - \lambda - 1/2 + s/2}{n(n+s)} \qquad (3.19)$$

tends to zero for any x as $n \to \infty$. From the same formula we see that all terms beginning from a certain one have the same sign. Hence the partial sums of the series grow faster as $x \to \infty$ than any finite power of x, so that condition (3.11) does not hold. This means that the second case is unsuitable, and we have

$$\lambda = \frac{s+1}{2} + p$$

Thus the only solution to Eq. (3.12) that satisfies our conditions is the polynomial

$$Q_p = a_0 F(-p;\ s+1;\ x) \qquad (3.20)$$

or in explicit form

$$Q_p = a_0 \left(1 - \frac{p}{1!} \frac{x}{s+1} + \frac{p(p-1)}{2!} \frac{x^2}{(s+1)(s+2)} + \cdots \right.$$

$$\left. \cdots + (-1)^p \frac{x^p}{(s+1)\ldots(s+p)}\right) \qquad (3.20^*)$$

If we put

$$a_0 = (s+1)\ldots(s+p) = \frac{\Gamma(s+p+1)}{\Gamma(s+1)} \qquad (3.21)$$

the corresponding functions Q_p, which we denote by $Q_p^s(x)$, are polynomials in s as well as in x:

$$Q_p^s(x) = \frac{\Gamma(s+p+1)}{\Gamma(s+1)} F(-p; s+1; x) \qquad (3.22)$$

or

$$Q_p^s(x) = (-1)^p \left(x^p - \frac{p}{1!}(s+p)x^{p-1} \right.$$
$$+ \frac{p(p-1)}{2!}(s+p)(s+p-1)x^{p-2}$$
$$\left. \cdots + (-1)^p(s+p)\cdots(s+1) \right) \qquad (3.22^*)$$

These are the *generalized Laguerre polynomials*; for $s = 0$ they are called simply *Laguerre polynomials*.

4. Some properties of generalized Laguerre polynomials

The generalized Laguerre polynomials, which are the solutions to the differential equation

$$x \frac{d^2Q_p^s}{dx^2} + (s+1-x)\frac{dQ_p^s}{dx} + pQ_p^s = 0 \qquad (4.1)$$

can be presented in the following form:

$$Q_p^s(x) = \frac{e^x}{x^s} \frac{d^p}{dx^p} e^{-x}x^{s+p} \qquad (4.2)$$

To prove this let us multiply Eq. (3.22*) by $x^s e^{-x}$ and write the product thus:

$$x^s e^{-x}Q_p^s(x) = \frac{d^p}{dx^p}\left(e^{-x}\right)x^{p+s} + p\frac{d^{p-1}}{dx^{p-1}}\left(e^{-x}\right)\frac{d}{dx}x^{p+s}$$
$$+ \frac{p(p-1)}{2!}\frac{d^{p-2}}{dx^{p-2}}(e^{-x})\frac{d^2}{dx^2}x^{p+s} + \cdots + e^{-x}\frac{d^p}{dx^p}x^{p+s}$$

According to Leibnitz's theorem for the nth derivative of the product of two functions,

$$x^s e^{-x}Q_p^s(x) = \frac{d^p}{dx^p} e^{-x}x^{s+p} \qquad (4.2^*)$$

which is what we set out to prove.

If we use the Cauchy formula, we can present this as

$$x^s e^{-x}Q_p^s(x) = \frac{p!}{2\pi i} \int \frac{e^{-z}z^{p+s}}{(z-x)^{p+1}} dz \qquad (4.3)$$

Next we introduce a new variable

$$t = \frac{z - x}{z}$$

and get

$$Q_p^s(x) = \frac{p!}{2\pi i} \int e^{-xt/(1-t)} \frac{1}{(1-t)^{s+1}} \frac{dt}{t^{p+1}} \tag{4.4}$$

But according to the Cauchy formula

$$Q_p^s(x) = \left(\frac{d^p}{dt^p} \frac{e^{-xt/(1-t)}}{(1-t)^{s+1}} \right)_{t=0} \tag{4.5}$$

Hence we obtain the Taylor series involving $Q_p^s(x)$:

$$(1-t)^{-s-1} e^{-xt/(1-t)} = \sum_{p=0}^{\infty} \frac{t^p}{p!} Q_p^s(x) \tag{4.6}$$

This is a convenient formula for deriving various relationships between generalized Laguerre polynomials. If we multiply it by $1 - t$, we get

$$(1-t)^{-s} e^{-xt/(1-t)} = \sum_{p=0}^{\infty} \frac{t^p}{p!} \left[Q_p^s(x) - p Q_{p-1}^s(x) \right] \tag{4.7}$$

On the other hand, changing s to $s - 1$ in (4.6), we get the same expression on the left-hand side. We then compare the coefficients in the two Taylor series and get

$$Q_p^{s-1}(x) = Q_p^s(x) - p Q_{p-1}^s(x) \tag{4.8}$$

This formula allows us to express generalized Laguerre poly-nomials with values of s that differ from one another by an integer, in terms of such polynomials with equal (the maximal of the two) values of s.

Let us differentiate both sides of (4.6) with respect to x. If we follow the same reasoning, we get

$$\frac{dQ_p^s(x)}{dx} = - p Q_{p-1}^{s+1}(x) \tag{4.9}$$

Now let us differentiate (4.7) with respect to t and express both sides of the results in terms of series. We then find that

$$sQ_p^s(x) - xQ_p^{s+1}(x) = Q_{p+1}^s(x) - (p+1) Q_p^s(x)$$

or, after we change s to $s - 1$,

$$xQ_p^s(x) = (p+s) Q_p^{s-1}(x) - Q_{p+1}^{s-1}(x) \tag{4.10}$$

Last, we use (4.8) and obtain the recursion relation that connects three successive polynomials of the same order, s:

$$(2p + s + 1 - x) Q_p^s (x) = Q_{p+1}^s (x) + p (p + s) Q_{p-1}^s (x) \quad (4.11)$$

It is also easy to find other relations by using (4.8), (4.9), and (4.10):

$$x \frac{dQ_p^s (x)}{dx} + sQ_p^s (x) = (p + s) Q_p^{s-1} (x) \quad (4.12)$$

$$x \frac{dQ_p^s (x)}{dx} + (s - x) Q_p^s (x) = Q_{p+1}^{s-1} (x) \quad (4.12^*)$$

Differentiating (4.12*) with respect to x and using (4.9), we arrive at Eq. (4.1) for $Q_p^s (x)$.

These formulas, in turn, make it possible to derive the following relations:

$$x \frac{dQ_{p-1}^s (x)}{dx} + (p + s - x) Q_{p-1}^s (x) = Q_p^s (x) \quad (4.13)$$

$$x \frac{dQ_p^s (x)}{dx} - pQ_p^s (x) = - p (p + s) Q_{p-1}^s (x) \quad (4.13^*)$$

which also lead to differential equation (4.1).

In the future we will have to evaluate integrals of type

$$J = \int_0^\infty x^s e^{-x} Q_p^s (x) f (x) dx \quad (4.14)$$

For this it is convenient to transform the integral using (4.2) and integrating by parts p times. We will have

$$J = \int_0^\infty \frac{d^p}{dx^p} (e^{-x} x^{p+s}) f (x) dx = (-1)^p \int_0^\infty e^{-x} x^{p+s} f^{(p)} (x) dx \quad (4.15)$$

Putting
$$f (x) = e^{(1-a) x}$$
we get

$$\int_0^\infty x^s e^{-ax} Q_p^s (x) dx = (a - 1)^p \int_0^\infty e^{-ax} x^{s+p} dx = \frac{(a - 1)^p}{a^{s+p+1}} \Gamma (s + p + 1) \quad (4.16)$$

A more general type of integral,

$$\int_0^\infty x^{s+r} e^{-ax} Q_p^s (x) dx \quad (4.17)$$

can be evaluated for integral values of r by differentiating (4.16) with respect to parameter a.

In (4.14) and (4.15) let us put $f(x) = x^r$, which yields

$$\int_0^\infty x^{s+r} e^{-x} Q_p^s(x)\, dx = (-1)^p r(r-1)\ldots(r-p+1)\Gamma(s+r+1) \quad (4.18)$$

If $f(x)$ is a polynomial of a degree lower than p, integral (4.14) vanishes. Using this property we can evaluate (4.14) when

$$f(x) = x^2 Q_p^s(x) = (-1)^p \left(x^{p+2} - p(s+p)x^{p+1} \right.$$
$$\left. + \frac{p(p-1)}{2}(s+p)(s+p-1)x^p + \ldots \right)$$

$$f(x) = x Q_p^s(x) = (-1)^p \left[x^{p+1} - p(s+p)x^p + \ldots \right]$$

$$f(x) = Q_p^s(x) = (-1)^p x^p + \ldots$$

$$f(x) = \frac{1}{x} Q_p^s(x) = \frac{\Gamma(s+p+1)}{\Gamma(s+1)} \frac{1}{x} + \ldots$$

$$f(x) = \frac{1}{x^2} Q_p^s(x) = \frac{\Gamma(s+p+1)}{\Gamma(s+1)} \left(\frac{1}{x^2} - \frac{p}{s+1} \frac{1}{x} + \ldots \right)$$

where the omitted terms are polynomials of a degree lower than p. This yields

$$\int_0^\infty e^{-x} x^{s+2} \left[Q_p^s(x) \right]^2 dx$$
$$= p!\Gamma(s+p+1) \left[6p^2 + 6p(s+1) + (s+1)(s+2) \right] \quad (4.19)$$

$$\int_0^\infty e^{-x} x^{s+1} \left[Q_p^s(x) \right]^2 dx = p!\Gamma(s+p+1)(2p+s+1) \quad (4.20)$$

$$\int_0^\infty e^{-x} x^s \left[Q_p^s(x) \right]^2 dx = p!\Gamma(s+p+1) \quad (4.21)$$

$$\int_0^\infty e^{-x} x^{s-1} \left[Q_p^s(x) \right]^2 dx = p!\Gamma(s+p+1)\frac{1}{s} \quad (4.22)$$

$$\int_0^\infty e^{-x} x^{s-2} \left[Q_p^s(x) \right]^2 dx = p!\Gamma(s+p+1)\frac{2p+s+1}{(s-1)s(s+1)} \quad (4.23)$$

Let us show that the polynomial $Q_p^s(x)$ has exactly p positive zeros, and that these zeros are real. If the number of zeros, say q, were less than p, then, denoting them by $\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_q$, we

could build the function

$$f(x) = (x - \alpha_1)(x - \alpha_2) \ldots (x - \alpha_q)$$

The product of this function by $Q_p^s(x)$ would then not change sign if x were to change from 0 to ∞, and (4.14) would not vanish. But this is impossible since $f(x)$ is a polynomial of a degree lower than p, and, according to (4.15), integral (4.14) must vanish. Hence the number of zeros cannot be less than p, and since it cannot be greater than p, it must be p.

5. Eigenvalues and eigenfunctions of the auxiliary problem

We return to the auxiliary problem of Section 3. The problem was to find the eigenfunctions and eigenvalues of the equation

$$-\frac{d}{dx}\left(x\,\frac{dy}{dx}\right) + \left(\frac{x}{4} + \frac{s^2}{4x}\right)y = \lambda y \tag{5.1}$$

We found the eigenvalues to be

$$\lambda = p + \frac{s+1}{2}, \qquad p = 0, 1, 2, \ldots \tag{5.2}$$

and the eigenfunctions were expressed in terms of the generalized Laguerre polynomials:

$$y_p(x) = c_p x^{s/2} e^{-x/2} Q_p^s(x) \tag{5.3}$$

The constant c_p we will find from the normalization condition

$$\int_0^\infty [y_p(x)]^2\,dx = 1$$

Recalling formula (4.21), we evaluate the integral and get

$$c_p = \frac{1}{[p!\Gamma(s+p+1)]^{1/2}} \tag{5.4}$$

Thus the functions

$$y_p(x) = \frac{1}{[p!\Gamma(s+p+1)]^{1/2}} x^{s/2} e^{-x/2} Q_p^s(x) \tag{5.5}$$

are orthonormalized:

$$\int_0^\infty y_p(x)\,y_{p'}(x)\,dx = \delta_{pp'} \tag{5.6}$$

and $\{y_p(x)\}$ is a complete set.

If we use (3.22), we can write

$$y_p(x) = \frac{1}{\Gamma(s+1)} \left(\frac{\Gamma(s+p+1)}{p!} \right)^{1/2} x^{s/2} e^{-x/2} F(-p; s+1; x) \quad (5.7)$$

It is sometimes convenient to use the normalized polynomials

$$Q_p^{*s}(x) = \frac{Q_p^s(x)}{[p!\Gamma(s+p+1)]^{1/2}} \quad (5.8)$$

which yields the following expression for $y_p(x)$:

$$y_p(x) = x^{s/2} e^{-x/2} Q_p^{*s}(x) \quad (5.9)$$

6. Energy levels and radial functions for the discrete hydrogen spectrum

Let us now solve our physical problem. First we will find the hydrogen energy levels. The energy parameter in atomic units, ε, was related to parameter λ of the auxiliary problem thus:

$$\frac{1}{(-2\varepsilon)^{1/2}} = \lambda \quad (6.1)$$

where λ was equal to

$$\lambda = p + \frac{s+1}{2}, \qquad p = 0, 1, 2, \ldots \quad (6.2)$$

Parameter s was connected with the azimuthal quantum number,

$$s = 2l + 1 \quad (6.3)$$

and p (integral values only) was equal to the number of nodes of the corresponding radial function, that is, according to the definition in Section 8, Chapter IV, to the radial quantum number, n_r. Hence λ is the integer

$$\lambda = n_r + l + 1 = n \quad (6.4)$$

which by definition is the principal quantum number.

The energy levels in atomic units are

$$\varepsilon_n = -\frac{1}{2n^2}, \qquad n = 1, 2, \ldots \quad (6.5)$$

Hence we see that they depend only on the value of the principal quantum number. This special property of the Coulomb field is of a fundamental nature. It is connected with the group of transformations under which the Schrödinger equation (in mo-

mentum space) for the hydrogen atom remains invariant. This group, which characterizes the increased symmetry of the hydrogen atom, coincides with the *rotational group* of a four-dimensional sphere. We will return to this point in Part IV.

In conventional units the energy levels of the hydrogen atom are

$$E = \frac{e^2}{a} \varepsilon_n = -\frac{2\pi R \hbar}{n^2} \tag{6.6}$$

where

$$R = \frac{me^4}{4\pi \hbar^3} = \frac{2\pi^2 me^4}{h^3} \tag{6.7}$$

is the *Rydberg constant*. Its value is

$$R = 3.2898421 \times 10^{15} \, s^{-1} \tag{6.8}$$

According to the remark made in Section 2, if we want to account for the finiteness of the mass of the nucleus, we must in all formulas [in (6.7) as well] change the electron mass to the reduced mass

$$m' = \frac{mM}{m + M}$$

The frequencies of the spectral lines are given by the Bohr frequency relation

$$\nu_{nn'} = \frac{\omega_{nn'}}{2\pi} = R \left(\frac{1}{n'^2} - \frac{1}{n^2} \right) \tag{6.9}$$

If $n' = 1$ and n has values 2, 3, ..., we have a series of lines called the *Lyman series*. If $n' = 2$ and $n = 3, 4, \ldots$ we have the *Balmer series*, and if $n' = 3$ and $n = 4, 5, \ldots$ we have the *Paschen series*.

Let us now express the radial functions in terms of the generalized Laguerre polynomials. Argument x in these polynomials is related to the distance in atomic units, r_1, in the following way:

$$x = \frac{2r_1}{n} \tag{6.10}$$

[see formulas (3.1), (3.2), and (6.4)]. We use (2.9), (5.9), and (6.2)-(6.4) to get

$$R_{nl}(r_1) = c_n \left(\frac{2r_1}{n} \right)^l e^{-r_1/n} Q_{n-l-1}^{*2l+1} \left(\frac{2r_1}{n} \right) \tag{6.11}$$

where c_n is a normalization factor determined from the condition

$$\int_0^\infty r_1^2 [R_{nl}(r_1)]^2 \, dr_1 = 1 \tag{6.12}$$

To evaluate the integral we introduce the variable x according to (6.10). Then

$$c_n^2 \left(\frac{n}{2}\right)^3 \int_0^\infty x^{2l+2} e^{-x} \left[Q_{n-l-1}^{*2l+1}(x)\right]^2 dx = 1$$

If we recall the relation between n and l, and p and s, we can express the above integral in terms of the ratio of (4.20) to (4.21). This yields

$$\int_0^\infty x^{s+1} e^{-x} \left[Q_p^{*s}(x)\right]^2 dx = 2p + s + 1 = 2n$$

whence

$$c_n^2 = \frac{4}{n^4}, \quad \text{or} \quad c_n = \frac{2}{n^2} \tag{6.13}$$

so that the normalized radial functions are

$$R_{nl}(r_1) = \frac{2}{n^2} \left(\frac{2r_1}{n}\right)^l e^{-r_1/n} Q_{n-l-1}^{*2l+1} \left(\frac{2r_1}{n}\right) \tag{6.14}$$

If we now use the expression for Q^* in terms of the confluent hypergeometric function and keep in mind that l is an integer, we get

$$R_{nl}(r_1) = \frac{2}{n^{l+2}} \left[(n-l)(n-l+1) \ldots (n+l)\right]^{1/2}$$

$$\times \frac{(2r_1)^l}{(2l+1)!} e^{-r_1/n} F\left(-n+l+1; 2l+2; \frac{2r_1}{n}\right) \tag{6.15}$$

where according to definition (3.16)

$$F\left(-n+l+1; 2l+2; \frac{2r_1}{n}\right)$$

$$= 1 - \frac{n-l-1}{(2l+2)1!} \frac{2r_1}{n} + \frac{(n-l-1)(n-l-2)}{(2l+2)(2l+3)2!} \left(\frac{2r_1}{n}\right)^2 - \ldots \tag{6.16}$$

This leads us directly to the asymptotic expansion of $R_{nl}(r_1)$ for very large values of n. The limit of (6.16) as $n \to \infty$ is

$$1 - \frac{2r_1}{(2l+2)1!} + \frac{(2r_1)^2}{(2l+2)(2l+3)2!} - \ldots$$

$$= (2l+1)! (2r_1)^{-l-1/2} J_{2l+1}\left(\sqrt{8r_1}\right) \tag{6.17}$$

where J_{2l+1} is the *Bessel function* of order $2l+1$. This yields

$$\lim_{n \to \infty} n^{3/2} R_{nl}(r_1) = \left(\frac{2}{r_1}\right)^{1/2} J_{2l+1}\left(\sqrt{8r_1}\right) \tag{6.18}$$

Functions (6.18), in contrast, belong to the continuous spectrum. The functions $R_{nl}(r_1)$ are eigenfunctions of the Hamiltonian and

possess the property of being orthogonal and normalized:

$$\int_0^\infty R_{nl}(r_1) R_{n'l}(r_1) r_1^2 \, dr_1 = \delta_{nn'} \qquad (6.19)$$

These functions, however, do not make up a complete set because the Hamiltonian has a continuous spectrum in addition to the discrete spectrum.

Finally, we write some simple radial functions $R_{nl}(r_1)$

$$R_{10}(r_1) = 2e^{-r_1} \qquad (6.20)$$

$$R_{20}(r_1) = \frac{1}{\sqrt{2}} e^{-r_1/2} \left(1 - \frac{r_1}{2}\right) \qquad (6.21)$$

$$R_{21}(r_1) = \frac{1}{\sqrt{24}} r_1 e^{-r_1/2} \qquad (6.21^*)$$

$$R_{30}(r_1) = \sqrt{\frac{4}{27}} e^{-r_1/3} \left[1 - \frac{2r_1}{3} + \frac{1}{6}\left(\frac{2r_1}{3}\right)^2\right] \qquad (6.22)$$

$$R_{31}(r_1) = \sqrt{\frac{8}{243}} e^{-r_1/3} \left[1 - \frac{1}{4}\left(\frac{2r_1}{3}\right)\right] \qquad (6.22^*)$$

$$R_{32}(r_1) = \frac{1}{\sqrt{2430}} e^{-r_1/3} \left(\frac{2r_1}{3}\right)^2 \qquad (6.22^{**})$$

7. Solution of the differential equation for the continuous spectrum in the form of a definite integral

We turn to the case of the continuous spectrum. In Eq. (2.10), which we write again,

$$\frac{d^2y}{dr_1^2} + \frac{1}{r_1}\frac{dy}{dr_1} + \left(2\varepsilon + \frac{2}{r_1} - \frac{s^2}{4r_1^2}\right)y = 0 \qquad (7.1)$$

parameter ε is a positive number. The variable

$$x_1 = r_1 (8\varepsilon)^{1/2} \qquad (7.2)$$

is real. We also put

$$\lambda_1 = (2\varepsilon)^{-1/2} \qquad (7.3)$$

Equation (7.1) then reads

$$x_1 \frac{d^2y}{dx_1^2} + \frac{dy}{dx_1} + \left(\frac{x_1}{4} + \lambda_1 - \frac{s^2}{4x_1}\right)y = 0 \qquad (7.4)$$

which differs from Eq. (3.3) in the sign of one of the terms. We can obtain this equation directly if in (3.3) we put

$$x = ix_1, \qquad \lambda = -i\lambda_1 \qquad (7.5)$$

This means that we can use the results of Section 3 and state that the solution to (7.4), finite at $x = 0$, is the function

$$y = e^{-ix_1/2} x_1^{s/2} Q(x_1) \tag{7.6}$$

with Q satisfying the differential equation

$$x_1 \frac{d^2 Q}{dx_1^2} + (s + 1 - ix_1) \frac{dQ}{dx_1} + \left[\lambda_1 - \frac{i}{2}(s+1)\right] Q = 0 \tag{7.7}$$

This implies that Q can be represented by a series:

$$Q = gF\left(\frac{s+1}{2} + i\lambda_1; \ s + 1; \ ix_1\right) \tag{7.8}$$

We will need the asymptotic expansions of Q and y for large values of x_1. The easiest way to obtain them is to express Q in the form of a definite integral. Equation (7.7) can be solved by using Laplace's method, which consists of the following. We seek the solution to (7.7) in the form of a contour integral in the complex z plane:

$$Q = \int e^{ix_1 z} f(z) \, dz \tag{7.9}$$

where $f(z)$ is still to be determined. Substituting (7.9) into (7.7) and differentiating under the integral sign, we get

$$x_1 \int e^{ix_1 z} (-z^2 + z) f(z) \, dz$$
$$+ \int e^{ix_1 z}\left[\lambda_1 + \frac{i}{2}(s+1)(2z-1)\right] f(z) \, dz = 0$$

To get rid of the factor x_1 in the first term we integrate by parts. This yields

$$\int z(1-z) f(z) \, d(-ie^{ix_1 z})$$
$$= -ie^{ix_1 z} z(1-z) f(z)\Big|_a^b + i \int e^{ix_1 z} \frac{d}{dz}[z(1-z) f(z)] \, dz$$

where the limits of integration are denoted by a and b.

If we now require that the first term on the right-hand side of the above expression vanish, that is

$$e^{ix_1 z} z(1-z) f(z)\Big|_a^b = 0 \tag{7.10}$$

the substitution of (7.9) into (7.7) results in

$$i \int e^{ix_1 z}\left(z(1-z) \frac{df}{dz} - \frac{s-1}{2}(1-2z) f(z) - i\lambda_1 f(z)\right) dz = 0 \tag{7.11}$$

We can satisfy this equation if we nullify the integrand, that is, if $f(z)$ satisfies the differential equation

$$\frac{f'(z)}{f(z)} = \frac{s-1}{2} \frac{1-2z}{z(1-z)} + \frac{i\lambda_1}{z(1-z)} \tag{7.12}$$

which we solve and get

$$\log f(z) = \frac{s-1}{2} \log[z(1-z)] + i\lambda_1 \log\frac{z}{1-z} + \log c \quad (7.13)$$

Whence

$$f(z) = cz^{i\lambda_1+(s-1)/2}(1-z)^{-i\lambda_1+(s-1)/2}$$

We see that the solution to (7.7) is

$$Q = c\int e^{ix_1z}z^{i\lambda_1+(s-1)/2}(1-z)^{-i\lambda_1+(s-1)/2}\,dz \quad (7.14)$$

only if we choose the contour of integration so that (7.10) is satisfied and we write it as

$$e^{ix_1z}z^{i\lambda_1+(s+1)/2}(1-z)^{-i\lambda_1+(s+1)/2}\Big|_a^b = 0 \quad (7.15)$$

The solution to (7.7) must be finite at $x = 0$. Such a solution can be found if we take the contour of integration that passes from $z = 0$ to $z = 1$ along the real axis. This contour satisfies condition (7.15) only if $s + 1 > 0$, which is always the case since we assume that $s \geqslant 0$. Substituting the limits of integration into (7.14), we get

$$Q = c\int_0^1 e^{ix_1z}z^{i\lambda_1+(s-1)/2}(1-z)^{-i\lambda_1+(s-1)/2}\,dz \quad (7.14^*)$$

To see whether the integral really does coincide with the series (7.8) we expand the exponential into a power series and then integrate termwise. If we use *Euler's integral* (or the *complete beta function*)

$$B(p, q) = \int_0^1 z^{p-1}(1-z)^{q-1}\,dz = \frac{\Gamma(p)\,\Gamma(q)}{\Gamma(p+q)} \quad (7.16)$$

we find that

$$Q = c\sum_{k=0}^{\infty}\frac{(ix_1)^k}{k!}\int_0^1 z^{i\lambda_1+k+(s-1)/2}(1-z)^{-i\lambda_1+(s-1)/2}\,dz$$

$$= c\sum_{k=0}^{\infty}\frac{(ix_1)^k}{k!}\frac{\Gamma\left(\frac{s+1}{2}+k+i\lambda_1\right)\Gamma\left(\frac{s+1}{2}-i\lambda_1\right)}{\Gamma(s+k+1)}$$

or

$$Q = c\frac{\Gamma\left(\frac{s+1}{2}+i\lambda_1\right)\Gamma\left(\frac{s+1}{2}-i\lambda_1\right)}{\Gamma(s+1)}F\left(\frac{s+1}{2}+i\lambda_1;\ s+1;\ ix_1\right)$$

$$(7.17)$$

which is what we set out to prove. By comparing (7.8) and (7.17) we see that the constants g and c of (7.8) and (7.14*) are connected in the following way:

$$g = c \frac{\Gamma\left(\frac{s+1}{2} + i\lambda_1\right)\Gamma\left(\frac{s+1}{2} - i\lambda_1\right)}{\Gamma(s+1)} \qquad (7.18)$$

We have thus found an expression for the confluent hypergeometric function:

$$F\left(\frac{s+1}{2} + i\lambda_1; s+1; ix_1\right) = \frac{\Gamma(s+1)}{\Gamma\left(\frac{s+1}{2} + i\lambda_1\right)\Gamma\left(\frac{s+1}{2} - i\lambda_1\right)}$$

$$\times \int_0^1 e^{ix_1 z} z^{i\lambda_1 + (s-1)/2} (1-z)^{-i\lambda_1 + (s-1)/2}\, dz \quad (7.19)$$

Since (7.8) and (7.14*) are solutions to Eq. (7.7) and are finite at $x_1 = 0$, we could derive (7.18) simply by comparing the two at $x_1 = 0$. If in (7.19) we put $z = 1 - z_1$, we get the relationship

$$F\left(\frac{s+1}{2} + i\lambda_1; s+1; ix_1\right) = e^{ix_1} F\left(\frac{s+1}{2} - i\lambda_1; s+1; -ix_1\right)$$

$$(7.20)$$

This means that function y, defined by (7.6), will be real only if g in (7.8) is real, which is what we will be assuming from now on.

8. Derivation of the asymptotic expression

To derive the asymptotic expression of Q for large positive values of x_1, using (7.14), we deform the contour of integration in (7.14) as follows. Instead of joining points 0 and 1 by a straight line we construct a broken line that passes from 0 to iA, then from iA to $iA + 1$, and finally from $iA + 1$ to 1 (A is any finite positive number). Since in the part of the complex z plane lying between the deformed contour of integration and the straight line from 0 to 1 the integrand in (7.14) is a holomorphic function, the deformation will not change the value of the integral. If we let A grow to infinity, the integral along the section from iA to $iA + 1$ will tend to zero because of the exponential $e^{-x_1 A}$ under the integral sign. In the limit we get

$$Q = \int_0^1 e^{ix_1 z} f(z)\, dz = \int_0^{i\infty} e^{ix_1 z} f(z)\, dz + \int_{1+i\infty}^1 e^{ix_1 z} f(z)\, dz \qquad (8.1)$$

where $f(z)$ is found by solving (7.13). In the first integral on the right-hand side we put

$$z = \zeta e^{i\pi/2} \tag{8.2}$$

whereas in the second we put

$$1 - z = \zeta e^{-i\pi/2} \tag{8.2*}$$

The limiting points for ζ in both integrals are 0 and ∞. We have

$$Q = c e^{-\pi\lambda_1/2 + i(s+1)\,\pi/4} \int_0^\infty e^{-x_1\zeta} \zeta^{(s-1)/2 + i\lambda_1} (1 - i\zeta)^{(s-1)/2 - i\lambda_1} d\zeta$$

$$+ e^{ix_1} c e^{-\pi\lambda_1/2 - i(s+1)\,\pi/4} \int_0^\infty e^{-x_1\zeta} \zeta^{(s-1)/2 - i\lambda_1} (1 + i\zeta)^{(s-1)/2 + i\lambda_1} d\zeta \tag{8.3}$$

Now let us introduce a new variable

$$t = \zeta x_1 \tag{8.4}$$

which yields

$$Q = c e^{i\,(s+1)\,\pi/4 - \pi\lambda_1/2} \frac{\Gamma\left(\dfrac{s+1}{2} + i\lambda_1\right)}{x_1^{(s+1)/2 + i\lambda_1}} J$$

$$+ e^{ix_1} c e^{-i\,(s+1)\,\pi/4 - \pi\lambda.2} \frac{\Gamma\left(\dfrac{s+1}{2} - i\lambda_1\right)}{x_1^{(s+1)/2 - i\lambda_1}} \bar{J} \tag{8.5}$$

where J denotes the integral

$$J = \frac{1}{\Gamma\left(\dfrac{s+1}{2} + i\lambda_1\right)} \int_0^\infty e^{-t} t^{(s-1)/2 + i\lambda_1} \left(1 - i\frac{t}{x_1}\right)^{(s-1)/2 - i\lambda_1} dt \tag{8.6}$$

and $\bar{J}$ the complex conjugate of J. The asymptotic expression for J can easily be obtained. For this we need only expand the integrand in inverse powers of x_1 and then integrate termwise. In view of the fact that the series

$$\left(1 - i\frac{t}{x_1}\right)^{(s-1)/2 - i\lambda_1} = 1 + \frac{i\lambda_1 - (s-1)/2}{1!} \frac{it}{x_1}$$

$$+ \frac{\left(i\lambda_1 - \dfrac{s-1}{2}\right)\left(i\lambda_1 - \dfrac{s-1}{2} + 1\right)}{2!} \left(\frac{it}{x_1}\right)^2 + \cdots \tag{8.7}$$

converges only if $|t| < |x_1|$, whereas with respect to t we integrate to infinity, the series obtained by termwise integration will

be a *divergent*, or *asymptotic, series*. Thus we have

$$J = F_{20}\left(i\lambda_1 + \frac{1}{2} - \frac{s}{2};\ i\lambda_1 + \frac{1}{2} + \frac{s}{2};\ \frac{i}{x_1}\right) \qquad (8.8)$$

where $F_{20}(\alpha;\ \beta;\ z)$ is a *formal power series*, that is

$$F_{20}(\alpha;\ \beta;\ z) = 1 + \frac{\alpha\beta}{1!}z + \frac{\alpha(\alpha + 1)\beta(\beta + 1)}{2!}z^2 + \ldots \qquad (8.9)$$

By (7.17) we can write our results as follows:

$$e^{-ix_1/2}F\left(\frac{s+1}{2} + i\lambda_1;\ s + 1;\ ix_1\right) = e^{ix_1/2}F\left(\frac{s+1}{2} - i\lambda;\ s + 1;\ -ix_1\right)$$

$$= \frac{\Gamma(s+1)}{\Gamma\left(\frac{s+1}{2} - i\lambda_1\right)}\ e^{i(s+1)\pi/4 - \pi\lambda_1/2}x_1^{-(s+1)/2 - i\lambda_1}e^{-ix_1/2}$$

$$\times F_{20}\left(i\lambda_1 + \frac{1-s}{2};\ i\lambda_1 + \frac{1+s}{2};\ \frac{i}{x_1}\right)$$

$$+ \frac{\Gamma(s+1)}{\Gamma\left(\frac{s+1}{2} + i\lambda_1\right)}\ e^{-i(s+1)\pi/4 - \pi\lambda_1/2}x_1^{-(s+1)/2 + i\lambda_1}e^{ix_1/2}$$

$$\times F_{20}\left(-i\lambda_1 + \frac{1-s}{2};\ -i\lambda_1 + \frac{1+s}{2};\ -\frac{i}{x_1}\right) \qquad (8.10)$$

where the equality is to be understood as asymptotic equality. The last formula holds not only for real values of λ_1 and x_1 but for their complex values as well, provided $-\pi/2 \leqslant \text{arc } x_1 \leqslant \pi/2$. For instance, if we put

$$\lambda_1 = i\lambda = i\left(\frac{s+1}{2} + p\right), \qquad x_1 = -ix$$

with p a positive integer, the second term in (8.10) will vanish because

$$\frac{1}{\Gamma(-p)} = 0$$

and (8.10) will give us not only asymptotic but exact equality:

$$e^{-x/2}F(-p;\ s + 1;\ x)$$

$$= \frac{(-1)^p\,\Gamma(s+1)}{\Gamma(s+p+1)}\ e^{-x/2}x^p F_{20}\left(-p - s;\ -p;\ -\frac{1}{x}\right) \qquad (8.11)$$

It is clear from a comparison of (8.11) with (3.11) and (3.22) that (8.11) gives polynomials $Q_p^s(x)$ in increasing powers of x (on the left) and diminishing powers of x (on the right).

9. Radial functions for the continuous hydrogen spectrum

On the basis of the results of the previous sections [see (2.9), (2.11), (7.2), (7.3), (7.6), and (7.8)] we can write an expression for the radial wave functions for the continuous spectrum of the hydrogen atom:

$$R_{\varepsilon l}(r_1) = a(\varepsilon)\, e^{-ir_1\sqrt{2\varepsilon}} \left(r_1\sqrt{8\varepsilon}\right)^l$$
$$\times F\left(l+1+\frac{i}{\sqrt{2\varepsilon}};\quad 2l+2;\ ir_1\sqrt{8\varepsilon}\right) \quad (9.1)$$

The functions are real only if $a(\varepsilon)$ is real. To obtain the asymptotic expression of the function for large values of r_1 we substitute into (8.10) the actual values of s, λ_1, x_1 and put

$$\frac{1}{\Gamma\left(l+1+\frac{i}{\sqrt{2\varepsilon}}\right)} = \frac{1}{\left|\Gamma\left(l+1+\frac{i}{\sqrt{2\varepsilon}}\right)\right|}\, e^{i\alpha} \quad (9.2)$$

Also, we replace the series F_{20} by their asymptotic value $F_{20} = 1$. We then get

$$R_{\varepsilon l}(r_1) \approx a(\varepsilon)\, \frac{(2l+1)!}{\left|\Gamma\left(l+1+\frac{i}{\sqrt{2\varepsilon}}\right)\right|}\, e^{-\pi/\sqrt{8\varepsilon}}\, \frac{1}{\sqrt{2\varepsilon}}$$
$$\times \frac{1}{r_1}\cos\left(r_1\sqrt{2\varepsilon} + \frac{1}{\sqrt{2\varepsilon}}\log\left(r_1\sqrt{8\varepsilon}\right) - (l+1)\frac{\pi}{2} + \alpha\right) \quad (9.3)$$

which we normalize thus:

$$\lim_{\Delta\varepsilon\to 0} \frac{1}{\Delta\varepsilon}\int_0^\infty r_1^2 \left|\int_\varepsilon^{\varepsilon+\Delta\varepsilon} R_{\varepsilon l}(r_1)\, d\varepsilon\right|^2 dr_1 = 1 \quad (9.4)$$

It is sometimes useful to introduce instead of ε a new parameter k:

$$\varepsilon = f(k) \quad (9.5)$$

where $f(k)$ is a monotonic function. We can then consider eigenfunction $R(k, r_1)$ normalized thus:

$$\lim_{\Delta k\to 0} \frac{1}{\Delta k}\int_0^\infty r_1^2 \left|\int_k^{k+\Delta k} R(k, r_1)\, dk\right|^2 dr_1 = 1 \quad (9.6)$$

Let us find the relationship between the functions with different normalizations. If $\Delta\varepsilon$ and Δk are considered positive,

$$\Delta\varepsilon = f'(k)\,\Delta k$$

$$\int\limits_{\varepsilon}^{\varepsilon+\Delta\varepsilon} R_{\varepsilon l}(r_1)\,d\varepsilon = f'(k) \int\limits_{k}^{k+\Delta k} R(k, r_1)\,dk$$

because Δk is infinitesimal. Substituting into (9.4) and comparing with (9.6), we get

$$R(k, r_1) = R_{\varepsilon l}(r_1)\left|\frac{d\varepsilon}{dk}\right|^{1/2} \tag{9.7}$$

We can transform (9.4) or (9.6) in the following way. Since the proper differentials belonging to different parts of the continuous spectrum are orthogonal to each other, we can write (9.6) as

$$\lim_{\Delta k \to 0} \frac{1}{\Delta k} \int\limits_{0}^{\infty} r_1^2 \left\{ \int\limits_{k}^{k+\Delta k} \overline{R(k', r_1)}\,dk' \int\limits_{k-\Delta_1 k}^{k+\Delta_1 k} R(k'', r_1)\,dk'' \right\} dr_1 = 1 \tag{9.8}$$

where $\Delta_1 k > \Delta k$. We can pass on to the limit for $\Delta k \to 0$ with $\Delta_1 k$ remaining finite. This yields

$$\int\limits_{0}^{\infty} r_1^2 \overline{R(k, r_1)} \int\limits_{k-\Delta_1 k}^{k+\Delta_1 k} R(k', r_1)\,dk'\,dr_1 = 1 \tag{9.9}$$

Now let $R^0(k, r_1)$ be the unnormalized functions and $c(k)$ the normalization factor, so that

$$R(k, r_1) = c(k)\,R^0(k, r_1) \tag{9.10}$$

Since $\Delta_1 k$ is infinitesimal, we can always take $c(k)$ outside the integral sign and thus obtain an equation for $c(k)$:

$$\frac{1}{|c(k)|^2} = \int\limits_{0}^{\infty} r_1^2 \overline{R^0(k, r_1)} \int\limits_{k-\Delta_1 k}^{k+\Delta_1 k} R^0(k', r_1)\,dk'\,dr_1 \tag{9.11}$$

(On the right the integral does not depend on $\Delta_1 k$, although it might seem to at first glance.)

These considerations apply not only to this particular example but to the general case of normalizing eigenfunctions in a continuous spectrum.

To evaluate integral (9.11) we must partition into two parts the domain of integration with respect to r_1: one interval is from zero to a value $r_1 = A$, and the other is from A to ∞. The integral over the finite interval (from 0 to A) tends to zero as $\Delta_1 k \to 0$.

What remains is the integral from A to ∞:

$$\frac{1}{|c(k)|^2} = \lim_{\Delta_1 k \to 0} \int_A^\infty r_1^2 \overline{R^0(k, r_1)} \int_{k-\Delta_1 k}^{k+\Delta_1 k} R^0(k', r_1) \, dk' \, dr_1 \quad (9.12)$$

The advantage of this formula is that by making A sufficiently large we can use the asymptotic expression for $R^0(k, r_1)$.

We put

$$k = (2\varepsilon)^{1/2} \quad (9.13)$$

and for $R^0(k, r_1)$ use the function having the asymptotic expression

$$R^0(k, r_1) = \frac{1}{r_1} \cos\left(kr_1 + \frac{1}{k}\log r_1 + \gamma\right) \quad (9.14)$$

where γ is a vector function of k that can be found by comparing (9.14) with (9.3). It can be shown that when evaluating (9.12) it is possible to consider $k^{-1}\log r_1 + \gamma$ under the integral sign as being negligible, since these terms are small compared to the leading term, kr_1. Keeping only this term, we get

$$\frac{1}{|c(k)|^2} = \lim_{\Delta_1 k \to 0} \int_A^\infty \cos kr_1 \int_{k-\Delta_1 k}^{k+\Delta_1 k} \cos k'r_1 \, dk' \, dr_1$$

$$= \lim_{\Delta_1 k \to 0} \int_A^\infty (1 + \cos 2kr_1) \frac{\sin(\Delta_1 kr_1)}{r_1} \, dr_1 \quad (9.15)$$

We prove that

$$\lim_{\Delta_1 k \to 0} \int_A^\infty \frac{\cos 2kr_1 \sin \Delta_1 kr_1}{r_1} \, dr_1 = 0 \quad (9.16)$$

Indeed, this is equal to

$$\lim_{\Delta_1 k \to 0} \left(\frac{1}{2} \int_A^\infty \frac{\sin(2k + \Delta_1 k)r_1}{r_1} \, dr_1 - \frac{1}{2} \int_A^\infty \frac{\sin(2k - \Delta_1 k)r_1}{r_1} \, dr_1 \right)$$

and the brackets contain the difference of two convergent integrals, which at $\Delta_1 k = 0$ coincide. What remains is

$$\frac{1}{|c(k)|^2} = \lim_{\Delta k \to 0} \int_A^\infty \frac{\sin(\Delta_1 kr_1)}{r_1} \, dr_1 \quad (9.17)$$

If we introduce a new variable, $t = r_1 \Delta_1 k$, we get

$$\frac{1}{|c(k)|^2} = \lim_{\Delta_1 k \to 0} \int_{A\Delta_1 k}^\infty \frac{\sin t}{t} \, dt = \int_0^\infty \frac{\sin t}{t} \, dt = \frac{\pi}{2} \quad (9.18)$$

so that we finally obtain

$$c(k) = (2/\pi)^{1/2} \qquad (9.19)$$

Hence the eigenfunctions normalized with respect to k will have the asymptotic expression

$$R(k, r_1) \approx \left(\frac{2}{\pi}\right)^{1/2} \frac{1}{r_1} \cos\left(kr_1 + \frac{1}{k}\log r_1 + \gamma\right) \qquad (9.20)$$

whereas those normalized with respect to ε will have, according to (9.7), the asymptotic expression

$$R_{\varepsilon l}(r_1) \approx \left(\frac{1}{2\varepsilon}\right)^{1/4}\left(\frac{2}{\pi}\right)^{1/2} \frac{1}{r_1} \cos\left(r_1\sqrt{2\varepsilon} + \frac{1}{\sqrt{2\varepsilon}}\log r_1 + \gamma\right) \quad (9.21)$$

Formula (9.3) now provides the following expression for the factor $a(\varepsilon)$ in (9.1):

$$a(\varepsilon) = \frac{1}{\pi^{1/2}}(8\varepsilon)^{1/4}\frac{e^{\pi/\sqrt{8\varepsilon}}}{(2l+1)!}\left|\Gamma\left(l+1+\frac{i}{\sqrt{2\varepsilon}}\right)\right| \qquad (9.22)$$

We can also express

$$\left|\Gamma\left(l+1+\frac{i}{\sqrt{2\varepsilon}}\right)\right|^2 = \Gamma\left(l+1+\frac{i}{\sqrt{2\varepsilon}}\right)\Gamma\left(l+1-\frac{i}{\sqrt{2\varepsilon}}\right)$$

in terms of elementary functions. If we use notation (7.3), multiply

$$\Gamma(l+1+i\lambda_1) = (l+i\lambda_1)\ldots(1+i\lambda_1)\Gamma(1+i\lambda_1)$$

and

$$\Gamma(l+1-i\lambda_1) = (l-i\lambda_1)\ldots(1-i\lambda_1)\Gamma(1-i\lambda_1)$$

and use the relationship

$$\Gamma(1+i\lambda_1)\Gamma(1-i\lambda_1) = \frac{\pi\lambda_1}{\sinh\pi\lambda_1} \qquad (9.23)$$

we find that

$$|\Gamma(l+1+i\lambda_1)|^2 = (1^2+\lambda_1^2)(2^2+\lambda_1^2)\ldots(l^2+\lambda_1^2)\frac{\pi\lambda_1}{\sinh\pi\lambda_1} \quad (9.24)$$

or

$$\left|\Gamma\left(l+1+\frac{i}{\sqrt{2\varepsilon}}\right)\right|^2$$
$$= \left(1+\frac{1}{2\varepsilon}\right)\left(4+\frac{1}{2\varepsilon}\right)\ldots\left(l^2+\frac{1}{2\varepsilon}\right)\frac{\pi/(2\varepsilon)^{1/2}}{\sinh[\pi/(2\varepsilon)^{1/2}]} \qquad (9.24^*)$$

Substitution of (9.24*) into (9.22) yields the final expression for $a(\varepsilon)$:

$$a(\varepsilon) = \frac{2\{[1+1/(2\varepsilon)]\ldots[l^2+1/(2\varepsilon)]\}^{1/2}}{\{1-\exp[-\pi(2/\varepsilon)^{1/2}]\}^{1/2}(2l+1)!} \qquad (9.25)$$

With this value of $a(\varepsilon)$ formula (9.1), namely

$$R_{\varepsilon l}(r_1) = a(\varepsilon) e^{-ir_1 \sqrt{2\varepsilon}} (r_1 \sqrt{8\varepsilon})^l$$
$$\times F\left(l + 1 + \frac{l}{\sqrt{2\varepsilon}};\ 2l + 2;\ ir_1 \sqrt{8\varepsilon}\right) \quad (9.1)$$

gives the normalized radial functions for the continuous hydrogen spectrum.

To end Section 9 we give the asymptotic expression for $R_{\varepsilon l}(r_1)$ as $\varepsilon \to 0$. By (6.17) and (6.18) we have

$$\lim_{\varepsilon \to 0} R_{\varepsilon l}(r_1) = \left(\frac{2}{r_1}\right)^{1/2} J_{2l+1}\left(\sqrt{8 r_1}\right) = \lim_{n \to \infty} n^{3/2} R_{nl}(r_1) \quad (9.26)$$

10. Intensities in the hydrogen spectrum

Knowing the radial function, we can calculate the intensities of spectral lines, which correspond to various transitions, and the intensities in the continuous spectrum. We recall the formulas for intensities which we derived in Section 9, Chapter IV. We also introduce atomic units and use the fact that in hydrogen the energy levels do not depend on the value of the azimuthal quantum number, l. Thus for transitions in the discrete spectrum

$$I(n'l;\ nl - 1) = (\varepsilon_{n'} - \varepsilon_n)^4 |r(n'l;\ nl - 1)|^2 l \quad (10.1)$$
$$I(n'l;\ nl + 1) = (\varepsilon_{n'} - \varepsilon_n)^4 |r(n'l;\ nl + 1)|^2 (l + 1) \quad (10.1^*)$$

and for transitions from the discrete spectrum into the continuous spectrum

$$I(n'l;\ \varepsilon)\Delta\varepsilon$$
$$= (\varepsilon_{n'} - \varepsilon)^4 [\ |r'(n'l;\ \varepsilon l - 1)|^2 l + |r'(n'l;\ \varepsilon l + 1)|^2 (l + 1)]\Delta\varepsilon \quad (10.2)$$

For arbitrary values of quantum numbers n', n, l the integers $r(n'l;\ nl \pm 1)$ are rather complicated. But for small values of n' the different $R_{n'l}$ are simple polynomials, and hence we can easily integrate with respect to r for arbitrary n or ε. We restrict ourselves to the Lyman $(n' = 1)$ and Balmer $(n' = 2)$ series.

At $n' = 1$ the only possible value of l' is zero, which means that we either evaluate

$$r(10;\ n1) = \int_0^\infty r_1^3 R_{10}(r_1) R_{n1}(r_1)\, dr_1 \quad (10.3)$$

or

$$r(10;\ \varepsilon 1) = \int_0^\infty r_1^3 R_{10}(r_1) R_{\varepsilon 1}(r_1)\, dr_1 \quad (10.3^*)$$

At $n' = 2$ there are two values of l', zero and unity. At $l' = 0$

$$r(20;\, n1) = \int_0^\infty r_1^3 R_{20}(r_1) R_{n1}(r_1)\, dr_1 \tag{10.4}$$

and at $l' = 1$

$$r(21;\, n0) = \int_0^\infty r_1^3 R_{21}(r_1) R_{n0}(r_1)\, dr_1 \tag{10.5}$$

$$r(21;\, n2) = \int_0^\infty r_1^3 R_{21}(r_1) R_{n2}(r_1)\, dr_1 \tag{10.6}$$

with similar integrals for the continuous spectrum. All the integrals can be evaluated in closed form with the help of (4.8) and (4.16) and their immediate generalizations for continuous-spectrum eigenfunctions. These formulas are

$$Q_p^s(x) = Q_p^{s+1}(x) - pQ_{p-1}^{s+1}(x) \tag{4.8}$$

$$\int_0^\infty x^s e^{-ax} Q_p^s(x)\, dx = \frac{(a-1)^p}{a^{s+p+1}} \Gamma(s+p+1) \tag{4.16}$$

To obtain this generalization let us express the generalized Laguerre polynomial Q_p^s in terms of the confluent hypergeometric function, F, using formula (3.22). We obtain

$$\begin{aligned}(s+1)&F(-p;\, s+1;\, x)\\ &= (s+1+p)\,F(-p;\, s+2;\, x) - pF(-p+1;\, s+2;\, x)\end{aligned} \tag{10.7}$$

$$\int_0^\infty e^{-ax} x^s F(-p;\, s+1;\, x)\, dx = \frac{\Gamma(s+1)}{a^{s+1}}\left(1 - \frac{1}{a}\right)^p \tag{10.8}$$

In the second expression we can introduce a new parameter b by changing a to a/b and x to bx. This yields

$$\int_0^\infty e^{-ax} x^s F(-p;\, s+1;\, bx)\, dx = \frac{\Gamma(s+1)}{a^{s+1}}\left(1 - \frac{b}{a}\right)^p \tag{10.9}$$

Our results are valid not only for integral values of p but for any fractional and complex values as well if constants a and b satisfy the condition $|a| > |b|$ and integrals (10.8) and (10.9) exist. To prove that (10.7) holds it is sufficient to compare the coefficients of powers of x on both sides of (10.7), and to prove (10.8) and (10.9) we only have to express F in the form of a series and integrate termwise.

We now turn to integral (10.3). Substituting the radial functions from (6.14) and (6.20), we get

$$r(10; \, nl)$$

$$= \frac{4}{3} n^{-5/2} \left(1 - \frac{1}{n^2}\right)^{1/2} \int_0^\infty r^4 e^{-(1 + 1/n)r} F\left(-n + 2; \, 4; \, \frac{2r}{n}\right) dr \quad (10.10)$$

Clearly, this is not an integral of type (10.9) since in (10.10) the power of r coincides with the second argument in F, whereas in (10.9) this argument exceeds the power of x by unity. To evaluate (10.10) we must either differentiate (10.9) with respect to a or, using (10.7), increase the second argument in F by unity and then use (10.9). In either case we get

$$\int_0^\infty r^4 e^{-(1 + 1/n)r} F\left(-n + 2; \, 4; \, \frac{2r}{n}\right) dr$$

$$= 12 \left(1 - \frac{1}{n}\right)^{n-3} \left(1 + \frac{1}{n}\right)^{-n-3} \quad (10.11)$$

Substitution into (10.10) yields

$$r(10; \, nl) = 16 n^{7/2} \frac{(n-1)^{n-5/2}}{(n+1)^{n+5/2}} \quad (10.12)$$

Integrals (10.4)-(10.6) are evaluated using the same method. After lengthy calculations we get

$$r(20; \, nl) = 2^8 \sqrt{2} \, n^{7/2} \left(n^2 - 1\right)^{1/2} \frac{(n-2)^{n-3}}{(n+2)^{n+3}} \quad (10.13)$$

$$r(21; \, n0) = \frac{2^8}{\sqrt{6}} n^{9/2} \frac{(n-2)^{n-3}}{(n+2)^{n+3}} \quad (10.14)$$

$$r(21; \, n2) = \frac{2^{10}}{\sqrt{6}} n^{9/2} \left(n^2 - 1\right)^{1/2} \frac{(n-2)^{n-7/2}}{(n+2)^{n+7/2}} \quad (10.15)$$

In the same way we can calculate the integrals for the continuous spectrum. For this we need not repeat all the steps but simply use (10.12)-(10.15). We reason as follows. If we compare (6.15) with (9.1) and (9.25), we see that the radial functions of the discrete and continuous spectra can be expressed in the following form:

$$R_{nl}(r_1) = \frac{2}{n^{3/2}} \Omega(n, \, l, \, r_1) \quad (10.16)$$

$$R_{\varepsilon l}(r_1) = \frac{2\Omega(1/(-2\varepsilon)^{1/2}, \, l, \, r_1)}{\{1 - \exp[-\pi(2/\varepsilon)^{1/2}]\}^{1/2}} \quad (10.17)$$

where in both cases Ω is the same analytic function of n, namely

$$\Omega\,(n,\,l,\,r) = \left[\left(1 - \tfrac{1}{n^2}\right)\left(1 - \tfrac{4}{n^2}\right) \cdots \left(1 - \tfrac{l^2}{n^2}\right)\right]^{1/2}$$
$$\times \frac{(2r)^l}{(2l+1)!}\, e^{-r/n} F\left(-n+l+1;\, 2l+2;\, \tfrac{2r}{n}\right) \quad (10.18)$$

On the other hand, integrals of type (10.9) are also analytic functions of their parameters. For instance, formula (10.11) holds for pure imaginary $n = i/(2\varepsilon)^{1/2}$. Hence, if we put

$$\tfrac{1}{2}\, n^{3/2} r\,(10;\, n1) = 8\left(1 - \tfrac{1}{n}\right)^{n-5/2}\left(1 + \tfrac{1}{n}\right)^{-n-5/2} = f\,(n) \quad (10.19)$$

then by (10.3), (10.3*), (10.16), and (10.17) we get

$$r'\,(10;\, \varepsilon 1) = \frac{2f\,(i/(2\varepsilon)^{1/2})}{\{1 - \exp\left[-\pi\,(2/\varepsilon)^{1/2}\right]\}^{1/2}} \quad (10.20)$$

or
$$r'\,(10;\, \varepsilon 1)$$

$$= \frac{16\,[1 + i\,(2\varepsilon)^{1/2}]^{i/\sqrt{2\varepsilon}-5/2}\,[1 - i\,(2\varepsilon)^{1/2}]^{-i/\sqrt{2\varepsilon}-5/2}}{\{1 - \exp\left[-\pi\,(2/\varepsilon)^{1/2}\right]\}^{1/2}} \quad (10.20^*)$$

We could obtain $r'\,(20;\, \varepsilon 1)$, $r'\,(21;\, \varepsilon 0)$, and $r'\,(21;\, \varepsilon 2)$ in a similar manner, but this can be done in a simpler way: by transferring to the continuous spectrum in the final formulas for intensity.

To find the intensity of transition between two energy levels we must sum (10.1) and (10.1*) over all possible values of l.

For the Lyman series the sum reduces to one term

$$I\,(1;\, n) = \left(\tfrac{1}{2} - \tfrac{1}{2n^2}\right)^4 |\,r\,(10;\, n1)\,|^2$$

so that by (10.12)

$$I\,(1;\, n) = \frac{2^4}{n}\,\frac{(n-1)^{2n-1}}{(n+1)^{2n+1}} \quad (10.21)$$

For the Balmer series we have

$$I\,(2;\, n) = \tfrac{1}{2^4}\left(\tfrac{1}{4} - \tfrac{1}{n^2}\right)^4$$
$$\times [\,|\,r\,(20;\, n1)\,|^2 + |\,r\,(21;\, n0)\,|^2 + 2\,|\,r\,(21;\, n2)\,|^2\,]$$

which according to (10.13)-(10.15) yields

$$I\,(2;\, n) = \frac{2^3}{n}\,\frac{(n-2)^{2n-3}}{(n+2)^{2n+3}}\,(5n^2 - 4)\,(3n^2 - 4) \quad (10.22)$$

If we want to find the intensities in the continuous spectrum, we must multiply (10.21) and (10.22) by the square of the ratio of

the factors of Ω in (10.16) and (10.17), namely, by

$$\frac{n^3}{1 - \exp\left[-\pi\,(2/\varepsilon)^{1/2}\right]}$$

after which we change n to $i/(2\varepsilon)^{1/2}$. The result reads

$$I\,(1;\,\varepsilon)\,\Delta\varepsilon$$

$$= \frac{16\left[1 + i\,(2\varepsilon)^{1/2}\right]^{2i/\sqrt{2\varepsilon}-1}\left[1 - i\,(2\varepsilon)^{1/2}\right]^{-2i/\sqrt{2\varepsilon}-1}\Delta\varepsilon}{1 - \exp\left[-\pi\,(2/\varepsilon)^{1/2}\right]} \tag{10.23}$$

$$I\,(2;\,\varepsilon)\,\Delta\varepsilon$$

$$= \frac{8\,(5 + 8\varepsilon)\,(3 + 8\varepsilon)\left[1 + 2i\,(2\varepsilon)^{1/2}\right]^{2i/\sqrt{2\varepsilon}-3}}{1 - \exp\left[-\pi\,(2/\varepsilon)^{1/2}\right]}$$

$$\times\left[1 - 2i\,(2\varepsilon)^{1/2}\right]^{-2i/\sqrt{2\varepsilon}-3}\Delta\varepsilon \tag{10.24}$$

To get rid of imaginary quantities we put in (10.23)

$$(2\varepsilon)^{1/2} = \tan\eta_1, \qquad 0 < \eta_1 < \pi/2 \tag{10.25}$$

and in (10.24)

$$2\,(2\varepsilon)^{1/2} = \tan\eta_2, \qquad 0 < \eta_2 < \pi/2 \tag{10.26}$$

Such substitutions lead to the following result:

$$I\,(1;\,\varepsilon)\,\Delta\varepsilon = \frac{16\exp\,(-4\eta_1\cot\eta_1)\tan\eta_1\,\Delta\eta_1}{1 - \exp\,(-2\pi\cot\eta_1)} \tag{10.27}$$

$$I\,(2;\,\varepsilon)\,\Delta\varepsilon$$

$$= \frac{2\exp\,(-8\eta_2\cot\eta_2)\,(1 + 4\cos^2\eta_2)\,(1 + 2\cos^2\eta_2)\tan\eta_2\,\Delta\eta_2}{1 - \exp\,(-4\pi\cot\eta_2)} \tag{10.28}$$

In conclusion we note that the intensities observed in experiments depend not only on the properties of separate atoms but on the number of atoms in the "initial" state, which will differ with experimental conditions. For this reason a comparison of our results with experience can be only indirect.

11. The Stark effect. General remarks

If an atom is placed inside an electric field, its energy levels and hence spectral lines, which correspond to transitions between the levels, are split, generally speaking, into several components. The splitting of spectral lines in an electric field is called the *Stark effect.* For the hydrogen atom this splitting is proportional to the field; for other atoms it is proportional to the square of the field. This difference can be explained by the following. When there is an electric field directed, say, along the z axis, the x component of angular momentum is a constant of the motion, which implies that the magnetic quantum number, m, still assumes a

definite value. In the case of a non-Coulomb field each energy level $E = E_{nl}$ will have corresponding to it only one eigenfunction with a definite value of m, whereas for a Coulomb field there will be several eigenfunctions. (They will differ by their azimuthal quantum number l.) In other words, the energy levels for a non-Coulomb field will be nondegenerate, and for a Coulomb field degenerate. This requires applying perturbation theory in different ways. Since the additional term in the potential energy is proportional to the z coordinate, the first-order correction to the unperturbed problem is zero for nondegenerate eigenvalues. Indeed, this correction is the diagonal matrix element, which vanishes owing to the selection rule for z. Hence for a non-Coulomb field it is the second-order correction that does not vanish, and this is proportional to the square of the external field. The situation changes for a Coulomb field (degenerate eigenvalues). The first order correction, which, we recall, is proportional to the field, is calculated by nullifying the determinant $D(\lambda)$ [Eq. (5.4), Chapter II], and it so happens that the correction is not zero.

We note that the state of an atom placed in an electric field is not stationary, strictly speaking. From the fact that for very great distances from the atom the electron's potential energy tends to $-\infty$ it becomes evident that there is a (nonzero) probability for the electron to separate itself from the atom. Thus perturbation theory provides an approximate solution to the Schrödinger equation in the sense of Section 1, Chapter II, and Section 8, Chapter III, that is, a quasi-stationary state.

In studying the Stark effect in hydrogen it is convenient to shift to parabolic coordinates instead of applying perturbation theory directly to the Schrödinger equation in spherical coordinates. This method has the advantage that in parabolic coordinates the perturbed equation allows for a separation of variables, and the problem reduces to solving equations whose unperturbed eigenvalues are nondegenerate. We will restrict ourselves to the first-order correction to energy levels.

12. The Schrödinger equation in parabolic coordinates

We write the eigenvalue equation for the Hamiltonian in atomic units:

$$\frac{1}{2} \nabla^2 \psi - \frac{1}{r} \psi - gz\psi = E\psi \tag{12.1}$$

The perturbation Hamiltonian is $-gz$, where g is the electric field in atomic units. If we denote the electric field in conventional units by D, then

$$D = \frac{e}{a^2} g = g \times 5.14 \times 10^9 \text{ V cm}^{-1} \tag{12.2}$$

This is why even in very strong fields of 10^5 or 10^6 V cm^{-1} parameter g will still be small.

Let us introduce parabolic coordinates

$$u = r + z, \quad v = r - z \tag{12.3}$$

The surfaces $u = $ constant and $v = $ constant represent an orthogonal system of paraboloids. This can be seen from the equations

$$\begin{aligned} x^2 + y^2 + 2uz &= u^2 \\ x^2 + y^2 - 2vz &= v^2 \end{aligned} \tag{12.4}$$

It is convenient to perform the transition to parabolic coordinates in three steps. First, we introduce the cylindrical coordinates z, ρ, φ. Then we put

$$z + i\rho = \frac{1}{2} (\xi + i\eta)^2 \tag{12.5}$$

Finally, we express ξ and η in terms of u and v. Separating the real and imaginary parts in (12.5), we get

$$z = \frac{1}{2} (\xi^2 - \eta^2) \tag{12.6}$$

$$\rho = \xi\eta \tag{12.7}$$

On the other hand, the modulus of (12.5) is

$$r = \frac{1}{2} (\xi^2 + \eta^2) \tag{12.8}$$

Whence

$$\xi^2 = r + z = u, \qquad \eta^2 = r - z = v \tag{12.9}$$

Next, we square the modulus of the differential of (12.5), express ξ and η in terms of u and v, and get

$$dz^2 + d\rho^2 = (\xi^2 + \eta^2)(d\xi^2 + d\eta^2)$$
$$= \frac{1}{4} (u + v) \left(\frac{1}{u} du^2 + \frac{1}{v} dv^2 \right)$$

so that the square of the element of arc length

$$ds^2 = dz^2 + d\rho^2 + \rho^2 d\varphi^2 \tag{12.10}$$

is

$$ds^2 = \frac{u+v}{4u} du^2 + \frac{u+v}{4v} dv^2 + uv \, d\varphi^2 \tag{12.11}$$

The square root of the product of the three terms on the right-hand side of (12.11) is the volume element:

$$d\tau = \frac{1}{4} (u + v) \, du \, dv \, d\varphi \tag{12.12}$$

What remains to be found is the expression for the Laplacian operator in parabolic coordinates. The simplest way to do this is to write the general expression for the Laplacian operator in terms of curvilinear orthogonal coordinates q_1, q_2, q_3:

$$\nabla^2\psi = \frac{1}{h_1 h_2 h_3}\left[\frac{\partial}{\partial q_1}\left(\frac{h_2 h_3}{h_1}\frac{\partial\psi}{\partial q_1}\right)\right.$$
$$\left. + \frac{\partial}{\partial q_2}\left(\frac{h_3 h_1}{h_2}\frac{\partial\psi}{\partial q_2}\right) + \frac{\partial}{\partial q_3}\left(\frac{h_1 h_2}{h_3}\frac{\partial\psi}{\partial q_3}\right)\right] \quad (12.13)$$

where h_1, h_2, h_3 are the square roots of the coefficients in the square of the element of arc length:

$$ds^2 = h_1^2\, dq_1^2 + h_2^2\, dq_2^2 + h_3^2\, dq_3^2 \quad (12.14)$$

Applying (12.13) to our case, we find that

$$\nabla^2\psi = \frac{4}{u+v}\left[\frac{\partial}{\partial u}\left(u\frac{\partial\psi}{\partial u}\right) + \frac{\partial}{\partial v}\left(v\frac{\partial\psi}{\partial v}\right) + \frac{u+v}{4uv}\frac{\partial^2\psi}{\partial\varphi^2}\right] \quad (12.15)$$

Now we substitute (12.15) into (12.1) and by (12.3) express r and z in terms of u and v. After multiplying by $(u+v)/2 = r$ and carrying all terms to the left-hand side we have

$$\frac{\partial}{\partial u}\left(u\frac{\partial\psi}{\partial u}\right) + \frac{\partial}{\partial v}\left(v\frac{\partial\psi}{\partial v}\right) + \frac{1}{4}\left(\frac{1}{u}+\frac{1}{v}\right)\frac{\partial^2\psi}{\partial\varphi^2}$$
$$+ \left[1 + \frac{1}{2}E(u+v) + \frac{g}{4}(u^2 - r^2)\right]\psi = 0 \quad (12.16)$$

Obviously the equation can be solved by a separation of variables. Indeed, if we put

$$\psi = U(u)\, V(v)\, e^{im\varphi} \quad (12.17)$$

Eq. (12.16) is satisfied if $U(u)$ and $V(v)$ are the solutions to the following equations:

$$\frac{d}{du}\left(u\frac{dU}{du}\right) + \left(a + \frac{1}{2}Eu - \frac{m^2}{4u} + \frac{g}{4}u^2\right)U = 0 \quad (12.18)$$

$$\frac{d}{dv}\left(v\frac{dV}{dv}\right) + \left(b + \frac{1}{2}Ev - \frac{m^2}{4v} - \frac{g}{4}v^2\right)V = 0 \quad (12.19)$$

where

$$a + b = 1 \quad (12.20)$$

Here m is simply the magnetic quantum number. Parameters a and b, connected by (12.20), are found from the condition that for all values of u and v from 0 to ∞ the solutions to Eqs. (12.18) and (12.19) are finite.

13. Splitting of energy levels in an electric field

Our aim is to find the corrections to the hydrogen energy levels for small values of the principal quantum number, n, assuming that g is very small.

Let us introduce the new variables

$$u_1 = u(-2E)^{1/2}, \qquad v_1 = v(-2E)^{1/2} \tag{13.1}$$

Since parameter E is negative [its approximate value is $E = -1/(2n^2)$], u_1 and v_1 are real and change from 0 to ∞. Equations (12.18) and (12.19) in the new variables read

$$\frac{d}{du_1}\left(u_1 \frac{dU}{du_1}\right) + \left(a_1 - \frac{1}{4}u_1 - \frac{m^2}{4u_1} + \frac{g_1}{4}u_1^2\right) U = 0 \tag{13.2}$$

$$\frac{d}{dv_1}\left(v_1 \frac{dV}{dv_1}\right) + \left(b_1 - \frac{1}{4}v_1 - \frac{m^2}{4v_1} - \frac{g_1}{4}v_1^2\right) V = 0 \tag{13.3}$$

where

$$g_1 = \frac{g}{(-2E)^{3/2}} \tag{13.4}$$

and where parameters a_1 and b_1 are

$$a_1 = \frac{a}{(-2E)^{1/2}}, \qquad b_1 = \frac{b}{(-2E)^{1/2}} \tag{13.5}$$

so that

$$a_1 + b_1 = (-2E)^{-1/2} \tag{13.6}$$

It follows from Eq. (13.4) that g_1 is a small constant only if the values of the principal quantum number, n, are small, a fact that we assumed at the start. If we consider g_1 known, Eqs. (13.2) and (13.3) are the eigenvalue equations for the hermitian operators with eigenvalues a_1 and b_1. If we find these eigenvalues, by using (13.6) we can find the energy eigenvalues, E.

In Eqs. (13.2) and (13.3) we will consider the terms that contain g_1 as the perturbation. The unperturbed equation is of the form

$$-\frac{d}{dx}\left(x \frac{dy}{dx}\right) + \left(\frac{x}{4} + \frac{m^2}{4x}\right) y = \lambda y \tag{13.7}$$

which we recognize to be the one we studied in Sections 3-5. Its eigenvalues are

$$\lambda = p + \frac{|m|+1}{2}, \qquad p = 0, 1, 2, \ldots \tag{13.8}$$

and its eigenfunctions are

$$y_p(x) = x^{|m|/2} e^{-x/2} Q_p^{*\,|m|}(x) \tag{13.9}$$

Hence in the zeroth-order approximation the eigenvalues and eigenfunctions of Eqs. (13.2) and (13.3) are

$$a_1^0 = n_1 + \frac{|m|+1}{2}, \qquad n_1 = 0, 1, 2, \ldots \qquad (13.10)$$

$$b_1^0 = n_2 + \frac{|m|+1}{2}, \qquad n_2 = 0, 1, 2, \ldots \qquad (13.11)$$

$$U_{n_1}^0 = y_{n_1}(u_1), \qquad V_{n_2}^0 = y_{n_2}(v_1) \qquad (13.12)$$

In the same approximation Eqs. (13.10) and (13.11) give

$$(-2E)^{-1/2} = a_1^0 + b_1^0 = n_1 + n_2 + |m| + 1 = n \qquad (13.13)$$

where n assumes the values $n = 1, 2, 3, \ldots$ and is simply the principal quantum number.

All the eigenvalues of the operator in (13.7) are non-degenerate. Therefore to obtain the first-order correction to (13.10) and (13.11) it is sufficient to find the diagonal matrix element of the perturbation, which is $- g_1 u_1^2/4$ in Eq. (13.2) and $g_1 v_1^2/4$ in Eq. (13.3). We have already evaluated the integral $\int_0^\infty x^2 [y_p(x)]^2 dx$, and according to (4.19) and (4.21)

$$\int\limits_0^\infty x^{s+2} e^{-x} [Q_p^{*s}(x)]^2 dx = 6p^2 + 6p(s+1) + (s+1)(s+2) \qquad (13.14)$$

Hence the eigenvalues in the first-order approximation are

$$a_1 = n_1 + \frac{|m|+1}{2} - \frac{g_1}{4} \left[6n_1^2 + 6n_1(|m|+1) \right.$$
$$\left. + (|m|+1)(|m|+2) \right] \qquad (13.15)$$

$$b_1 = n_2 + \frac{|m|+1}{2} + \frac{g_1}{4} \left[6n_2^2 + 6n_2(|m|+1) \right.$$
$$\left. + (|m|+1)(|m|+2) \right] \qquad (13.16)$$

Their sum is

$$a_1 + b_1 = n + \frac{3}{2} g_1 n (n_2 - n_1) \qquad (13.17)$$

where n is determined by (13.13). Then owing to (13.4) and (13.6) we have the approximate equality

$$(-2E)^{-1/2} = n + \frac{3}{2} gn(n_2 - n_1)(-2E)^{-3/2} \qquad (13.18)$$

To simplify matters let us change $(-2E)^{1/2}$ in the correction term to its approximate value $1/n$. We have

$$(-2E)^{-1/2} = n + \frac{3}{2} gn^4 (n_2 - n_1) \qquad (13.19)$$

Solving this equation for E, we get in the same approximation

$$E = -\frac{1}{2n^2} + \frac{3}{2} gn (n_2 - n_1) \qquad (13.20)$$

Thus the energy levels depend, in fact, only on two quantum numbers: the principal quantum number, n, and the difference $n_2 - n_1$. For a given value of n this difference can assume all the values (integral, of course) from $-(n-1)$ to $(n-1)$. The pattern of the splitting of an energy state term in an external electric field can be found by taking specific values of $n_2 - n_1$ one after another. Formula (13.20) agrees with experiments.

We have seen that the state of the electron of the unperturbed hydrogen atom can be described either by specifying the quantum numbers n, l, m (in spherical coordinates) or by specifying the quantum numbers n_1, n_2, m (in parabolic coordinates). Since the eigenfunctions ψ_{nlm} do not coincide with $\psi'_{n_1 n_2 m}$ (the first are linear combinations of the second with the same values of n and m, and vice versa), the states nlm and $n_1 n_2 m$ differ. But then in what state is the electron of the unperturbed hydrogen atom? We can judge the state of the atom by measuring its energy E, which for an unperturbed hydrogen atom depends only on n. This means that in a state with a definite energy only this quantum number has a definite value, whereas the other quantum numbers (say, n_1, n_2, m) remain not only unknown but physically undefined. The most we can say is that the wave function of such an atom is a linear combination of the eigenfunctions with the fixed value of n and various values of the other quantum numbers (in fact, this is a combination of the functions ψ_{nlm} as well as the functions $\psi'_{n_1 n_2 m}$). To specify the value of $n_1 - n_2$ we must act on the atom in a definite way. Namely, we must place it in an electric field. When this is done and the difference $n_1 - n_2$ is measured, the state of the atom becomes known with greater accuracy. We can then say that here the wave function is expressed in terms of the eigenfunctions with definite values of n and $n_1 - n_2$. Hence it is the physical interaction with the atom that enables us to distinguish between states rather than mathematical arbitrariness (choice of coordinates or eigenfunctions).

14. Scattering of α-particles.
Statement of the problem

Let us imagine that there is a coordinate system with the heavy nucleus of an atom at its origin. A plane wave representing an α-particle with definite momentum $p_z = p > 0$ and energy E falls onto this nucleus from the negative half of the z axis (from the left). The wave undergoes diffraction: the outgoing wave from

the scattering centre interferes in the right half of the z axis ($z > 0$) with the incident wave. This means that the α-particle is "partially" deflected by the scattering centre and "partially" passes the scattering centre undisturbed (this must be understood as a linear superposition of states). For the sake of clarity we can visualize a whole plane of scattering centres and a beam of α-particles falling onto it in the direction of positive values of z. To the left of this plane the *flux* of α-particles is entirely along the z axis in the positive direction (that is, to the right). To the right of this plane it divides into two parts: the passing flux of α-particles and the deflected flux.

What we want is to describe this phenomenon by using the wave function ψ and to find the ratio of the flux of scattered α-particles to the incident flux as a function of the deflection angle.

We denote the charge and mass of the α-particle by $2e$ and m_α, the charge and mass of the nucleus (the target) by Ze and M, and put

$$\mu = \frac{m_\alpha M}{m_\alpha + M} \tag{14.1}$$

When an α-particle hits an atom of a high-Z element, Coulomb repulsion between the two plays the main role, whereas the action of the electron shell on the α-particle is negligible. We can therefore assume the potential energy to be

$$U(r) = \frac{2Ze^2}{r} \tag{14.2}$$

The Schrödinger equation of our problem is

$$-\frac{\hbar^2}{2\mu} \nabla^2 \psi + \frac{2Ze^2}{r} \psi - i\hbar \frac{\partial \psi}{\partial t} = 0 \tag{14.3}$$

and we will consider states with a definite energy equal to

$$E = \frac{p^2}{2\mu} \tag{14.4}$$

where p is the momentum of the α-particle at infinity. Let us also assume that the angular momentum of the α-particle about the z axis is zero, so that the wave function does not depend on angle φ.

At negative values of z and large values of r the wave function by definition must be a plane wave

$$\psi = e^{i(zp - Et)/\hbar} \tag{14.5}$$

with $-\infty < z < 0$ and $r \to \infty$. Aside from this, to make our solution unique we specify that no ingoing waves from infinity are present.

Since the energy has a definite value, we can put

$$\psi = \psi^0 e^{-iEt/\hbar} \tag{14.6}$$

where ψ^0 does not depend on t and satisfies the equation

$$-\frac{\hbar^2}{2\mu}\nabla^2\psi^0 + \frac{2Ze^2}{r}\psi^0 = E\psi^0 \tag{14.7}$$

Before solving this equation let us simplify its coefficients. This is done by introducing a new unit of length,

$$r_0 = \frac{\hbar^2}{2Ze^2\mu} \tag{14.8}$$

The new coordinates are

$$x' = \frac{x}{r_0}, \qquad y' = \frac{y}{r_0}, \qquad z' = \frac{z}{r_0}, \qquad r' = \frac{r}{r_0} \tag{14.9}$$

We also put

$$E = \varepsilon\,\frac{2Ze^2}{r_0} \tag{14.10}$$

Equation (14.7) then reads

$$-\frac{1}{2}\nabla^2_{r'}\psi^0 + \frac{1}{r'}\psi^0 = \varepsilon\psi^{0'} \tag{14.11}$$

and condition (14.5) takes the form

$$\psi^0 \sim \exp\left[i\,(2\varepsilon)^{1/2}\,z'\right] \tag{14.12}$$

at $-\infty < z' < 0$ and $r' \to \infty$.

15. Solution of equations

Since according to the conditions of the problem the z axis plays a special role, we introduce parabolic coordinates:

$$u = r' + z', \qquad v = r' - z' \tag{15.1}$$

Next we use the results of Section 12, specifically Eq. (12.16), in which we (a) change the term $+1$ in the brackets to -1 (because in our case the Coulomb energy has the opposite sign), (b) put $g = 0$, and (c) keep in mind that ψ^0 does not depend on angle φ. We get

$$\frac{\partial}{\partial u}\left(u\,\frac{\partial\psi^0}{\partial u}\right) + \frac{\partial}{\partial v}\left(v\,\frac{\partial\psi^0}{\partial v}\right) + \left[-1 + \frac{1}{2}\,\varepsilon\,(u+v)\right]\psi^0 = 0 \tag{15.2}$$

The condition at infinity, (14.12), is written in parabolic coordinates as

$$\psi^0 \sim \exp\left[i\,(2\varepsilon)^{1/2}\,(u-v)/2\right] \tag{15.3}$$

as $v \to \infty$ and for all values of u. We can satisfy this condition only if

$$\psi^0 = \exp\left[i\,(\varepsilon/2)^{1/2}u\right]V \tag{15.4}$$

where V does not depend on u but satisfies the asymptotic condition.

$$V \sim \exp\left[-i\,(\varepsilon/2)^{1/2}v\right] \quad \text{at} \quad v \to \infty \tag{15.5}$$

If we now substitute (15.4) into (15.2), we see that the former is a solution only if V satisfies the equation

$$\frac{d}{dv}\left(v\,\frac{dV}{dv}\right) + \left[i\left(\frac{\varepsilon}{2}\right)^{1/2} - 1 + \frac{1}{2}\,\varepsilon v\right]V = 0 \tag{15.6}$$

Since $\varepsilon > 0$, we can put

$$(2\varepsilon)^{1/2}v = v_1 \tag{15.7}$$

after which we obtain

$$\frac{d}{dv_1}\left(v_1\,\frac{dV}{dv_1}\right) + \left[\frac{i}{2} - \frac{1}{(2\varepsilon)^{1/2}} + \frac{1}{4}\,v_1\right]V = 0 \tag{15.8}$$

This equation coincides with the one we studied in detail in Sections 7 and 8:

$$\frac{d}{dx_1}\left(x_1\,\frac{dy}{dx_1}\right) + \left(\frac{x_1}{4} + \lambda_1 - \frac{s^2}{4x_1}\right)y = 0 \tag{7.4}$$

In our case

$$\lambda_1 = \frac{i}{2} - \frac{1}{(2\varepsilon)^{1/2}}, \qquad s = 0, \quad x_1 = v_1 \tag{15.9}$$

If for the sake of convenience we put

$$b = \frac{1}{(2\varepsilon)^{1/2}}, \qquad \lambda_1 = \frac{i}{2} - b \tag{15.10}$$

by (7.6) and (7.8) we can write the solution to Eq. (15.8), finite at $v_1 = 0$, as

$$V = ce^{-iv_1/2}F\left(-ib;\ 1;\ iv_1\right)$$
$$= ce^{-iv_1/2}\left[1 + \frac{(-ib)}{(1!)^2}\,iv_1 + \frac{(-ib)\,(-ib+1)}{(2!)^2}\,(iv_1)^2 + \dots\right] \tag{15.11}$$

Constant c can be found from condition (15.5), that is

$$V \sim e^{-iv_1/2} \quad \text{as} \quad v_1 \to \infty \tag{15.12}$$

This requires that we use the asymptotic expansion for the confluent hypergeometric function, F, derived in Section 8 [see (8.10)]. For the specific values of the parameters we have

$$V = \frac{ce^{\pi b/2}}{\Gamma(1 + ib)}\,e^{-i\,(v_1/2 - b\log v_1)}\,F_{20}\left(-ib;\ -ib;\ i/v_1\right)$$
$$- \frac{cbe^{\pi b/2}}{v_1\Gamma(1 - ib)}\,e^{i\,(v_1/2 - b\log v_1)}\,F_{20}\left(1 + ib;\ 1 + ib;\ -i/v_1\right) \tag{15.13}$$

where F_{20} are formal power series built according to (8.9). Condition (15.12) holds approximately if we put

$$c = e^{-\pi b/2}\Gamma\,(1 + ib) \tag{15.14}$$

Let us now turn to coordinates r' and z' and build the function ψ^0. On the basis of (15.1), (15.4), and (15.7) we get the following expressions for ψ^0. For small values of $v = r' - z'$

$$\psi^0 = e^{-\pi b/2}\Gamma\,(1 + ib)\,e^{iz'/b}\Big[1 + \frac{(-ib)}{(1\,!)^2}\,\frac{i\,(r'-z')}{b}$$
$$+ \frac{(-ib)\,(-b+1)}{(2\,!)^2}\Big(\frac{i\,(r'-z')}{b}\Big)^2 + \ldots\Big] \tag{15.15}$$

and for large values of $v = r' - z'$

$$\psi^0 = e^{iz'/b}e^{ib\,\log\,[(r'-z')/b]}\Big(1 + \frac{(-ib)^2}{1}\,\frac{ib}{r'-z'} + \ldots\Big)$$
$$- \frac{\Gamma\,(1+ib)}{\Gamma\,(1-ib)}\,\frac{b^2}{r'-z'}\,e^{ir'/b}e^{-ib\,\log\,[(r'-z')/b]}\,(1 + \ldots) \tag{15.16}$$

Note that parameter b is proportional to the wavelength.

16. The Rutherford scattering law

Formula (15.16) provides an overall solution for our problem. At large distances from the atom and not too close to the z axis (at least at a distance of several wavelengths from this axis) the wave function consists of two terms. The first term, being approximately

$$\psi_1^0 = e^{iz'/b + ib\,\log\,[(r'-z')/b]} \tag{16.1}$$

represents the passing plane wave corresponding to the undeflected beam of α-particles. The second term, approximately being

$$\psi_2^0 = - b^2\,\frac{\Gamma\,(1+ib)}{\Gamma\,(1-ib)}\,\frac{1}{r'-z'}\,e^{ir'/b-ib\,\log\,[(r'-z')/b]} \tag{16.2}$$

is the scattered spherical wave corresponding to the deflected particles.

To estimate the relative intensity of the deflected and undeflected particle fluxes, let us build, using (16.1) and then (16.2), two quantities that are proportional to the current densities. We recall formulas (2.10) and (2.11), Chapter III, and get

$$\mathbf{S}_1 = \mathrm{grad}'\Big(\frac{z'}{b} + b\,\log\frac{r'-z'}{b}\Big) \tag{16.3}$$

$$\mathbf{S}_2 = \frac{b^4}{(r'-z')^2}\,\mathrm{grad}'\Big(\frac{r'}{b} - b\,\log\frac{r'-z'}{b}\Big) \tag{16.4}$$

Obviously, vector S_1 at great distances is directed along the z axis, and S_2 along the radius vector away from the scattering centre.

If we denote the number of particles crossing an area $d\sigma$ per unit time by $n\,d\sigma$ and the number of particles scattered into the area that subtends a solid angle $d\omega$ at the origin (the target) per unit time by $\Phi\,d\omega$, then by using (16.3) and (16.4) we find

$$\Phi = nr_0^2 \left(\frac{r'}{r' - z'} \right)^2 b^4 \tag{16.5}$$

If we denote the *scattering angle* by θ, then

$$z' = r' \cos \theta \tag{16.6}$$

With this in mind we can write (16.5) in the following form:

$$\Phi = \frac{nr_0^2 b^4}{4 \sin^4 (\theta/2)} \tag{16.7}$$

or, if we replace r_0 and b by their values given by (14.8), (14.10) and (15.10),

$$\Phi = n \left(\frac{Ze^2}{2E} \right)^2 \frac{1}{\sin^4 (\theta/2)} \tag{16.8}$$

Ernest Rutherford derived this formula on the basis of classical mechanical notions and verified it in experiments. The law of inverse proportionality to the fourth power of the sine of the scattering half-angle was fully confirmed.

17. The virial theorem in classical and in quantum mechanics

Classical mechanics contains a theorem that deals with the motion of a system of particles in a finite region of space. During this motion the particles' coordinates and, naturally, their velocities remain finite. The *virial theorem* states that the time average value of kinetic energy T is related to the *virial*, which is a certain linear function of the derivatives, with respect to the rectangular coordinates, of potential energy U. The coefficients of the derivatives are proportional to the corresponding coordinates. For one mass point the virial is

$$V = x \frac{\partial U}{\partial x} + y \frac{\partial U}{\partial y} + z \frac{\partial U}{\partial z} \tag{17.1}$$

with U the potential energy. If we recall that the equations of motion are

$$m\ddot{x} = -\frac{\partial U}{\partial x}, \qquad m\ddot{y} = -\frac{\partial U}{\partial y}, \qquad m\ddot{z} = -\frac{\partial U}{\partial z} \tag{17.2}$$

we see that

$$V = -m(x\ddot{x} + y\ddot{y} + z\ddot{z}) \tag{17.3}$$

On the other hand, obviously,

$$\frac{d}{dt}(x^2 + y^2 + z^2) = 2(x\dot{x} + y\dot{y} + z\dot{z}) \tag{17.4}$$

$$\frac{d^2}{dt^2}(x^2 + y^2 + z^2) = 2(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) + 2(x\ddot{x} + y\ddot{y} + z\ddot{z}) \tag{17.5}$$

or

$$\frac{d^2}{dt^2}(x^2 + y^2 + z^2) = \frac{4}{m}T - \frac{2}{m}V \tag{17.6}$$

It is easy to see that the time average value of the left-hand side of (17.6) vanishes since it is equal to the difference of the values of (17.4) at two distant moments of time, divided by the time interval.

Thus

$$2\langle T\rangle = \langle V\rangle \tag{17.7}$$

This is the essence of the virial theorem of classical mechanics for the case of one mass point. The theorem can be generalized for a system of mass points.

If the potential energy is a homogeneous function of degree ρ in the (Cartesian) coordinates, the definition of the virial, (17.1), gives

$$V = \rho U \tag{17.8}$$

and hence

$$2\langle T\rangle = \rho\langle U\rangle \tag{17.9}$$

Now we take up the case involving quantum mechanics. The time average of a classical quantity has corresponding to it in quantum mechanics the mathematical expectation of the quantum analogue of this quantity in a state with definite energy. Let us show that, if we relate classical and quantum quantities in such a way, there is a condition in quantum mechanics similar to the virial theorem of classical mechanics.

The Schrödinger equation for the motion of a mass point (as well as for the motion of a system of mass points) can be obtained from the variational method. This method states that for actual motion

$$\delta I = 0 \tag{17.10}$$

where the *action integral, I*, is

$$I = \int \bar{\psi}(T + U - E)\psi \, d\tau \tag{17.11}$$

with T the kinetic-energy operator, U the potential energy operator, and E the energy parameter (total energy).

We will keep to the case of one mass point. Then

$$T\psi = -\frac{\hbar^2}{2m}\nabla^2\psi \qquad (17.12)$$

where ∇^2 is the Laplacian operator.

If ψ is normalized so that

$$\int \bar\psi\psi \, d\tau = 1 \qquad (17.13)$$

then

$$T_0 = \int \bar\psi T\psi \, d\tau \qquad (17.14)$$

is the mathematical expectation of kinetic energy and

$$U_0 = \int \bar\psi U\psi \, d\tau \qquad (17.15)$$

is the mathematical expectation of potential energy.

After these preliminary remarks we can formulate the virial theorem as follows.

If the wave function, ψ, belongs to the discrete spectrum [10] and the potential energy is a homogeneous function of degree ρ in the coordinates, that is, if

$$U(\lambda\mathbf{r}) = \lambda^\rho U(\mathbf{r}) \qquad (17.16)$$

we come to the following relationship:

$$2T_0 = \rho U_0 \qquad (17.17)$$

that is, the double mathematical expectation of kinetic energy is equal to ρ times the mathematical expectation of potential energy.

To prove the theorem we first change $\mathbf{r}$ in $\psi(\mathbf{r})$ to a proportional quantity $(\mathbf{r} \rightarrow \lambda\mathbf{r})$ and consider the function

$$\psi^*(\mathbf{r}, \lambda) = \lambda^{3/2}\psi(\lambda\mathbf{r}) \qquad (17.18)$$

which according to (17.13) is normalized to unity. Next we substitute ψ^* into action integral (17.11) and denote by T_0^* and U_0^* the mathematical expectations of the kinetic and potential energies in the state described by ψ^*. When x, y, z change to λx, λy, λz (that is, with the change of scale), T changes to $\lambda^2 T$ (normalization for ψ^* coincides with normalization for ψ). This means that

$$T_0^* = \int \bar\psi^* T\psi^* \, d\tau = \lambda^2 T_0 \qquad (17.19)$$

---

[10] This corresponds to a motion in a finite region of space and with a finite velocity.

and also that

$$U_0^* = \int \bar{\psi}^* U \psi^* \, d\tau = \int \overline{\psi(\mathbf{r})} \, U\left(\frac{\mathbf{r}}{\lambda}\right) \psi(\mathbf{r}) \, d\tau \qquad (17.20)$$

which owing to the homogeneity of U yields

$$U_0^* = \lambda^{-\rho} U_0 \qquad (17.21)$$

The action integral is then

$$I = \lambda^2 T_0 + \lambda^{-\rho} U_0 - E \qquad (17.22)$$

Finally, nullifying its variation in parameter λ, we get

$$\delta I = \left(2\lambda T_0 - \rho\lambda^{-\rho-1} U_0\right)\delta\lambda = 0 \qquad (17.23)$$

But the solution of the variational problem is found at $\lambda = 1$. Hence

$$2T_0 = \rho U_0 \qquad (17.24)$$

which is what we set out to prove.

In the general case of an arbitrary potential energy we have

$$2T_0 = -\left(\frac{\partial U_0^*}{\partial \lambda}\right)_{\lambda=1} = \mathsf{M.E.}\left(x\frac{\partial U}{\partial z} + y\frac{\partial U}{\partial y} + z\frac{\partial U}{\partial z}\right) = \mathsf{M.E.}\,V \tag{17.25}$$

We can apply the same reasoning to the case of a system of particles if operators T and U are homogeneous functions defined by (17.19) and (17.21).

The relationship expressing the virial theorem is satisfied not only by the exact solution but by approximate solutions as well, provided that these are obtained by using the variational method, which allows for variation of scale. Variation of scale means that the wave function can be transformed according to (17.18), with an appropriate generalization for many particles. After this transformation λ can be found by the variational method. For instance, the method of self-consistent field (see Part IV), obtained by the variational method, allows for variation of scale.

18. Some remarks concerning the superposition principle and the probabilistic interpretation of the wave function

In Sections 14-16 we considered the problem of a particle colliding with the heavy nucleus of an atom. This problem is of special interest because it provides a graphic demonstration of the physical meaning of the concept of a wave function.

We know that the wave function $\psi(x, y, z, t)$ serves to describe the state of one particle. This state can be such that a given physical quantity does not have a definite value. For instance, let

the wave function of a free particle be

$$\psi = (c_1 e^{ixp/\hbar} + c_2 e^{iyp/\hbar}) \exp\left(-\frac{i}{\hbar}\frac{p^2}{2m} t\right) \tag{18.1}$$

In this state the particle's energy is

$$E = \frac{p^2}{2m} \tag{18.2}$$

But we cannot specify the direction of its motion. If we were to run a large number of tests that made it possible to determine the direction of motion (say, with the help of a specially oriented aperture), we would find that there is a *probability* of the particle moving along the x axis with a velocity $v = p/m$. There would also be a probability of the particle moving along the y axis with the same velocity. In other words, in the state described by wave function (18.1) the particle has a *potential possibility* of being observed moving in one or the other direction. Obviously, this is a state different from the one in which the particle moves along the bisector of the angle between the x and y axes. In the latter case a particle having a velocity corresponding to the value of energy (18.2) would be in the state

$$\psi = c_3 e^{-iEt/\hbar} \exp\left(\frac{i}{\hbar} \frac{x+y}{\sqrt{2}} p\right) \tag{18.3}$$

which is completely different from (18.1). The possibility of the existence of states in which a given quantity does not have a definite value and which are linear combinations of states with definite values of this quantity (the *superposition principle*) is most characteristic of quantum mechanics, and herein lies the fundamental difference between quantum and classical mechanics. The language of classical mechanics cannot describe such a "mixed" state of one particle. The superposition principle is necessary, however, if we want to derive, starting from a general principle, the dual nature of light and matter, which manifest themselves as waves and as particles.

In the early history of wave mechanics the wave function was interpreted as a certain wave in physical space, a wave that was connected with a collection of particles capable of diffraction (de Broglie's wave). The superposition principle rested on this assumption. For instance, in (18.1) the term

$$c_1 e^{ixp/\hbar} \exp\left(-\frac{i}{\hbar}\frac{p^2}{2m} t\right)$$

corresponded to a wave representing a flux of particles moving along the x axis with velocity $v = p/m$, whereas the second term represented a flux of particles moving along the y axis. The whole wave function was, so to say, the superposition of these two fluxes,

and their ratio was equal to the ratio of the amplitudes' square moduli, $|c_1|^2/|c_2|^2$.

Such an interpretation of the wave function, thanks to de Broglie, was pictorial. Strictly speaking, however, it was not a precise interpretation because the two fluxes can interfere with each other and superposition can be understood as a linear combination of the wave functions representing the fluxes. More than that, we know that a wave function describes the state of only one particle (and not a flux) and must be interpreted on the basis of the potential possibility of certain results being obtained in experiments (measurements) with the particle. We have mentioned this fact in Section 6, Chapter IV, Part I. There we also noted that a quantum object cannot be an element of a statistical ensemble. For this reason there is no purpose in introducing "ensembles" of such objects, as is done in statistical physics. If we want to assume that the concept of probability requires the existence of an ensemble, this can only be an ensemble of the results of a definite set of experiments.

To return to the wave function (18.1). If we consider it from the angle of potential possibility, we can build formulas for the probabilities of finding a particle moving along the x and y axes, provided that initially the particle was in state (18.1). The probabilities will then be proportional to the squares of the amplitudes moduli, and their ratio will be equal to $|c_1|^2/|c_2|^2$. This result coincides with what we get when we consider de Broglie's waves as referring to fluxes of particles. But we should not understand the word "fluxes" too literally. We must allow for the possibility of mutual destruction of the waves that represent the fluxes as the result of interference of the waves.

Part III

PAULI'S THEORY OF THE ELECTRON

1. The electron angular momentum

In Section 7, Chapter II, we considered the operators of angular momentum, which we built using the position and momentum operators:

$$m_x = yp_z - zp_y$$
$$m_y = zp_x - xp_z$$
$$m_z = xp_y - yp_x \qquad (1.1)$$

These operators can represent the angular momentum of a mass point with three degrees of freedom, that is, moving in space. The behaviour of an electron in a magnetic field and the properties of systems of many electrons (the electron shell of an atom, for instance) show that the electron possesses an *internal* degree of freedom connected with its intrinsic angular momentum, which does not depend on the motion in space. This internal degree of freedom (or the corresponding intrinsic angular momentum) of the electron is called the *electron spin*.

The properties of the electron's intrinsic angular momentum (*spin*) can be studied by using the commutation relations between the components of the conventional (orbital) angular momentum

$$m_y m_z - m_z m_y = i\hbar m_x$$
$$m_z m_x - m_x m_z = i\hbar m_y$$
$$m_x m_y - m_y m_x = i\hbar m_z \qquad (1.2)$$

We assume here that *spin angular momentum* satisfies the same relations. We also introduce the hypothesis that the operators for each of the spin components possess only two eigenvalues that differ only in their signs.

The operators of spin angular momentum components can be written as

$$(m_x)_{\text{spin}} = \frac{\hbar}{2}\,\sigma_x, \quad (m_y)_{\text{spin}} = \frac{\hbar}{2}\,\sigma_y, \quad (m_z)_{\text{spin}} = \frac{\hbar}{2}\,\sigma_z \qquad (1.3)$$

where σ_x, σ_y, σ_z are hermitian operators with two nondegenerate eigenvalues, $+1$ and -1. The $\frac{1}{2}$-factor appears because operators (1.3) must satisfy (1.2).

When calculating spin in units of $\hbar$ (and not in units of $\hbar/2$), it is convenient to use the operators

$$s_x = \frac{1}{2}\,\sigma_x, \qquad s_y = \frac{1}{2}\,\sigma_y, \qquad s_z = \frac{1}{2}\,\sigma_z$$

instead of σ_x, σ_y, σ_z. We will use these new operators in Part IV to study the many-electron problem.

The commutation relations for σ_x, σ_y, σ_z are

$$\sigma_y\sigma_z - \sigma_z\sigma_y = 2i\sigma_x$$
$$\sigma_z\sigma_x - \sigma_x\sigma_z = 2i\sigma_y$$
$$\sigma_x\sigma_y - \sigma_y\sigma_x = 2i\sigma_z \tag{1.4}$$

At the same time there is the following identity:

$$\sigma_x^2\sigma_z - \sigma_z\sigma_x^2 = \sigma_x\left(\sigma_x\sigma_z - \sigma_z\sigma_x\right) + \left(\sigma_x\sigma_z - \sigma_z\sigma_x\right)\sigma_x$$

which together with (1.4) yields

$$\sigma_x^2\sigma_z - \sigma_z\sigma_x^2 = -2i\left(\sigma_x\sigma_y + \sigma_y\sigma_x\right)$$

According to our hypothesis, the eigenvalues of σ_x are ±1, so that σ_x^2 has only one eigenvalue $\sigma_x^2 = 1$. This means that σ_x^2 is not an operator but simply a number, which commutes with any operator, σ_z inclusive. Hence the right-hand side of the last equality vanishes. Together with similar equalities for other components this yields

$$\sigma_y\sigma_z + \sigma_z\sigma_y = 0$$
$$\sigma_z\sigma_x + \sigma_x\sigma_z = 0$$
$$\sigma_x\sigma_y + \sigma_y\sigma_x = 0 \tag{1.5}$$

which are called the *anticommutation relations* (the operators in this case are said to *anticommute*). If we compare (1.4) with (1.5), we can see that

$$\sigma_y\sigma_z = -\sigma_z\sigma_y = i\sigma_x$$
$$\sigma_z\sigma_x = -\sigma_x\sigma_z = i\sigma_y$$
$$\sigma_x\sigma_y = -\sigma_y\sigma_x = i\sigma_z \tag{1.6}$$

Also

$$\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = 1 \tag{1.7}$$

We can consider σ_x, σ_y, σ_z either as 2×2 matrices or as operators that act on functions of a certain new variable σ (in addition

to coordinates) that assumes only two values, $\sigma = \pm 1$, say. Denoting a function of this kind by $\psi(r, \sigma)$, where r represents all three position coordinates, we can satisfy relations (1.6) and (1.7) if, for instance, we put

$$\sigma_x \psi(r, \sigma) = \psi(r, -\sigma)$$
$$\sigma_y \psi(r, \sigma) = -i\sigma\psi(r, -\sigma)$$
$$\sigma_z \psi(r, \sigma) = \sigma\psi(r, \sigma) \tag{1.8}$$

If we write ψ as a *two component column function*

$$\psi = \begin{pmatrix} \xi \\ \eta \end{pmatrix} \tag{1.9}$$

where $\xi = \psi(r, +1)$ and $\eta = \psi(r, -1)$, then the previous formulas are written

$$\sigma_x \begin{pmatrix} \xi \\ \eta \end{pmatrix} = \begin{pmatrix} \eta \\ \xi \end{pmatrix}, \quad \sigma_y \begin{pmatrix} \xi \\ \eta \end{pmatrix} = \begin{pmatrix} -i\eta \\ i\xi \end{pmatrix}, \quad \sigma_z \begin{pmatrix} \xi \\ \eta \end{pmatrix} = \begin{pmatrix} \xi \\ -\eta \end{pmatrix} \tag{1.10}$$

Hence the operators can be represented in matrix form as

$$\sigma_x = \sigma_1, \qquad \sigma_y = \sigma_2, \qquad \sigma_z = \sigma_3 \tag{1.11}$$

where

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{1.12}$$

These are the *Pauli matrices*. The choice of matrices for operators σ_x, σ_y, σ_z is determined to within a canonical transformation (which corresponds to a linear combination of ξ and η). Whence, if we assume, as is customary in scientific literature, that σ_1, σ_2, σ_3 are matrices with numbers as elements, (1.12), and preserve the physical meaning of σ_x, σ_y, σ_z that follows from (1.8), then equalities (1.11) are not the only possible ones and can be replaced by equivalent equalities.

The operators σ_x, σ_y, σ_z, which satisfy (1.6) and (1.7), are of a vector nature in the sense that, if we introduce three linear forms

$$\sigma'_x = l_1\sigma_x + m_1\sigma_y + n_1\sigma_z$$
$$\sigma'_y = l_2\sigma_x + m_2\sigma_y + n_2\sigma_z$$
$$\sigma'_z = l_3\sigma_x + m_3\sigma_y + n_3\sigma_z \tag{1.13}$$

where l_k, m_k, n_k ($k = 1, 2, 3$) are the cosines of the angles between two rectangular coordinate systems, the new operators, σ'_x, σ'_y, will have the same properties (1.6) and (1.7) as the old operators, σ_x, σ_y, σ_z. It then follows that if we consider the three quantities to be the components of a vector, the eigenvalues of the projection

of this vector on any direction are ± 1. Note that when we assume operators $\sigma_x, \sigma_y, \sigma_z$ to be matrices (1.12), as was done in (1.11), then (1.13) are the most general relationships for 2×2 matrices that satisfy (1.6) and (1.7).

The three matrices (1.12) together with the unit matrix form a complete set in the sense that an arbitrary 2×2 matrix can be expressed in terms of a linear combination of these four matrices with coefficients that are complex numbers. If the matrix under consideration is hermitian, the coefficients are real.

Now we turn to the problem of *total angular momentum operators*. If we take the operators of spin angular momentum in the form (1.3), the operators of *total angular momentum* are obviously

$$\mathscr{M}_x = m_x + \frac{\hbar}{2}\sigma_x, \quad \mathscr{M}_y = m_y + \frac{\hbar}{2}\sigma_y, \quad \mathscr{M}_z = m_z + \frac{\hbar}{2}\sigma_z \quad (1.14)$$

Owing to the general property of angular momentum, operators (1.14) must satisfy the same commutation relations as the operators of orbital and spin angular momenta. This results in the following relations similar to (1.2):

$$\mathscr{M}_y\mathscr{M}_z - \mathscr{M}_z\mathscr{M}_y = i\hbar\mathscr{M}_x$$
$$\mathscr{M}_z\mathscr{M}_x - \mathscr{M}_x\mathscr{M}_z = i\hbar\mathscr{M}_y$$
$$\mathscr{M}_x\mathscr{M}_y - \mathscr{M}_y\mathscr{M}_x = i\hbar\mathscr{M}_z \quad (1.15)$$

These commutation relations can be verified directly by using relations (1.2) and (1.4) for m_x, m_y, m_z and $\sigma_x, \sigma_y, \sigma_z$ respectively and remembering that orbital angular momentum commutes with spin angular momentum.

We can use the components of the spin and orbital angular momenta to build a bilinear form that commutes with each component of the total angular momentum. Indeed, let us put

$$\mathscr{M} = \sigma_x m_x + \sigma_y m_y + \sigma_z m_z + \hbar \quad (1.16)$$

or, which is the same,

$$\mathscr{M} = \sigma_x\mathscr{M}_x + \sigma_y\mathscr{M}_y + \sigma_z\mathscr{M}_z - \frac{1}{2}\hbar \quad (1.16^*)$$

According to the properties of m_x, m_y, m_z,

$$\mathscr{M}m_x - m_x\mathscr{M} = -i\hbar(\sigma_y m_z - \sigma_z m_y)$$

whereas according to the properties of $\sigma_x, \sigma_y, \sigma_z$

$$\frac{\hbar}{2}(\mathscr{M}\sigma_x - \sigma_x\mathscr{M}) = i\hbar(\sigma_y m_z - \sigma_z m_y)$$

If we then add the two equalities, the right-hand side vanishes. The expression on the left together with the similar relations for

the other two components yields

$$\mathcal{M}\mathcal{M}_x - \mathcal{M}_x\mathcal{M} = 0$$
$$\mathcal{M}\mathcal{M}_y - \mathcal{M}_y\mathcal{M} = 0$$
$$\mathcal{M}\mathcal{M}_z - \mathcal{M}_z\mathcal{M} = 0 \qquad (1.17)$$

Let us see how operator $\mathcal{M}$ is connected with the operator of orbital angular momentum in Schrödinger's theory. We find that

$$\mathcal{M}^2 - \mathcal{M}\hbar = m_x^2 + m_y^2 + m_z^2 \qquad (1.18)$$

On the other hand, we can square operators corresponding to the orbital angular momentum components, (1.14). We then see that

$$\mathcal{M}_x^2 + \mathcal{M}_y^2 + \mathcal{M}_z^2 = \mathcal{M}^2 - \frac{1}{4}\hbar^2 \qquad (1.19)$$

Thus the square of vector $(\mathcal{M}_x, \mathcal{M}_y, \mathcal{M}_z)$ differs from the square of the scalar operator $\mathcal{M}$ only by the term $\hbar^2/4$. We will call $\mathcal{M}$ the *operator of the spin-orbit scalar of angular momentum.*

The right-hand side of (1.18) is an operator in Schrödinger's theory and does not contain the Pauli matrices. Its eigenvalues are $\hbar^2 l(l+1)$, where l is a positive integer or zero. If we denote the eigenvalues of $\mathcal{M}$ by $\hbar k$, we have

$$k(k-1) = l(l+1) \qquad (1.20)$$

which for a given value of l yields

$$k = -l \quad \text{or} \quad k = l+1 \qquad (1.21)$$

But k cannot be zero. Indeed, from formula (1.19) it follows that the mathematical expectation of $\mathcal{M}^2$ exceeds $\hbar^2/4$ in any state. Hence the eigenvalue equation

$$\mathcal{M}\psi = k\hbar\psi \qquad (1.22)$$

cannot have a zero eigenvalue. This in turn means that at $l = 0$ the only possible value of k is $k = 1$, and at $l \geq 1$ there are two values of k, given by formula (1.21).

2. The operators of total angular momentum in spherical coordinates

When we studied the problem of motion due to a central field in Schrödinger's theory (Chapter IV, Part II), we found the expressions for the operators of orbital angular momentum in spherical coordinates r, θ, φ. These coordinates are connected with rectangular coordinates x, y, z via the following formulas:

$$x = r \sin\theta \cos\varphi, \qquad y = r \sin\theta \sin\varphi, \qquad z = r \cos\theta \qquad (2.1)$$

If we put

$$p_r = -i\hbar \frac{\partial}{\partial r}, \qquad p_\theta = -i\hbar \frac{\partial}{\partial \theta}, \qquad p_\varphi = -i\hbar \frac{\partial}{\partial \varphi} \qquad (2.2)$$

then according to formulas (3.2), Chapter IV, Part II, we have

$$m_x = -\sin\varphi\, p_\theta - \cot\theta \cos\varphi\, p_\varphi$$
$$m_y = \cos\varphi\, p_\theta - \cot\theta \sin\varphi\, p_\varphi$$
$$m_z = p_\varphi \qquad (2.3)$$

The simultaneous eigenfunctions of the square of angular momentum and the projection of angular momentum on the z axis must in Schrödinger's theory satisfy the following equations:

$$\mathbf{m}^2 \psi = -\hbar^2 \left[\frac{1}{\sin\theta} \frac{\partial}{\partial \theta} \left(\sin\theta \frac{\partial \psi}{\partial \theta} \right) + \frac{1}{\sin^2\theta} \frac{\partial^2 \psi}{\partial \varphi^2} \right]$$
$$= \hbar^2 l(l+1)\, \psi \qquad (2.4)$$

$$m_z \psi = -i\hbar \frac{\partial \psi}{\partial \varphi} = m\psi \qquad (2.5)$$

and must be single-valued on the surface of a sphere.

Now we consider the total angular momentum, which includes spin. Let us reintroduce the operators $\mathcal{M}_z$ and $\mathcal{M}$ of the previous section [see (1.14) and (1.16)]. In terms of p_θ and p_φ these operators are

$$\mathcal{M}_z = p_\varphi + \frac{\hbar}{2} \sigma_z \qquad (2.6)$$

$$\mathcal{M} = (-\sigma_x \sin\varphi + \sigma_y \cos\varphi)\, p_\theta$$
$$+ (-\sigma_x \cot\theta \cos\varphi - \sigma_y \cot\theta \sin\varphi + \sigma_z)\, p_\varphi + \hbar \qquad (2.7)$$

We ask for the simultaneous eigenfunctions of $\mathcal{M}_z$ and $\mathcal{M}$. These functions must satisfy the equations

$$\mathcal{M}_z \psi = \hbar \left(m + \frac{1}{2} \right) \psi \qquad (2.8)$$

$$\mathcal{M} \psi = \hbar k \psi \qquad (2.9)$$

and must be single-valued on the surface of a sphere.

As to the eigenvalues of the two operators, we have already stated that for a given nonzero value of l the quantum number k can assume the values $k = -l$ and $k = l + 1$, whereas at $l = 0$ the only possible value of k is unity. The quantum number m is the same as in Schrödinger's theory. It can assume all integral values from $m = -l$ to $m = +l$.

To simplify (2.6) and (2.7) let us use the canonical transformation

$$L' = SLS^+, \qquad \psi' = S\psi \qquad (2.10)$$

where

$$S = \cos\frac{\varphi}{2} + i\sigma_z \sin\frac{\varphi}{2}, \qquad S^+ = \cos\frac{\varphi}{2} - i\sigma_z \sin\frac{\varphi}{2} \qquad (2.11)$$

If prior to transformation

$$\sigma_x = \sigma_1, \qquad \sigma_y = \sigma_2, \qquad \sigma_z = \sigma_3 \qquad (2.12)$$

$$\mathscr{M}_z = -i\hbar\frac{\partial}{\partial\varphi} + \frac{\hbar}{2}\sigma_3 \qquad (2.13)$$

then after transformation

$$\sigma_x' = \sigma_1\cos\varphi - \sigma_2\sin\varphi, \quad \sigma_y' = \sigma_2\sin\varphi + \sigma_2\cos\varphi, \quad \sigma_z' = \sigma_3 \qquad (2.14)$$

$$\mathscr{M}_z' = -i\hbar\frac{\partial}{\partial\varphi} \qquad (2.15)$$

Changing the operators σ_x, σ_y, σ_z in (2.7) according to (2.14) to σ_x', σ_y', σ_z' and p_φ to

$$p_\varphi' = Sp_\varphi S^+ = p_\varphi - \frac{\hbar}{2}\sigma_3 \qquad (2.16)$$

we find the transformed operator $\mathscr{M}$ in the form

$$\mathscr{M}' = -\sigma_1\cot\theta\, p_\varphi + \sigma_2\left(p_\theta - \frac{i\hbar}{2}\cot\theta\right) + \sigma_3 p_\varphi + \frac{\hbar}{2} \qquad (2.17)$$

where p_φ and p_θ are understood to be the operators (2.2).

To further simplify $\mathscr{M}$ we apply to it two transformations one after another. First, we put

$$\mathscr{M}'' = T\mathscr{M}'T^+, \qquad \psi'' = T\psi' \qquad (2.18)$$

where

$$T = \cos\frac{\theta}{2} + i\sigma_2\sin\frac{\theta}{2}, \qquad T^+ = \cos\frac{\theta}{2} - i\sigma_2\sin\frac{\theta}{2} \qquad (2.19)$$

We get

$$\mathscr{M}'' = -\frac{\sigma_1}{\sin\theta}p_\varphi + \sigma_2\left(p_\theta - \frac{i\hbar}{2}\cot\theta\right) \qquad (2.20)$$

Second, we put

$$\mathscr{M}^* = (\sin\theta)^{1/2}\mathscr{M}''(\sin\theta)^{-1/2}, \qquad \psi^* = (\sin\theta)^{1/2}\psi'' \qquad (2.21)$$

The transformed operator $\mathscr{M}^*$ now assumes the form

$$\mathscr{M}^* = -\frac{\sigma_1}{\sin\theta}p_\varphi + \sigma_2 p_\theta \qquad (2.22)$$

which is much simpler than the initial form, (2.7). Operator $\mathscr{M}_z'$, on the other hand, has not changed after transformations (2.18) and (2.21), so that we have

$$\mathscr{M}_z^* = p_\varphi \qquad (2.23)$$

If we consider ψ to be a two component wave function on the surface of a sphere [see (1.9)],

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} \tag{2.24}$$

and if we put

$$|\psi|^2 = |\psi_1|^2 + |\psi_2|^2 \tag{2.25}$$

we can take the normalization condition to be

$$\frac{1}{4\pi} \int\limits_0^\pi \int\limits_0^{2\pi} |\psi|^2 \sin\theta \, d\theta \, d\varphi = 1 \tag{2.26}$$

The same will hold for ψ' and ψ''. But for ψ^*, which differs from ψ'' by the factor $(\sin\theta)^{1/2}$, the normalization condition is

$$\frac{1}{4\pi} \int\limits_0^\pi \int\limits_0^{2\pi} |\psi^*|^2 \, d\theta \, d\varphi = 1 \tag{2.27}$$

Only the original wave function is required to be single-valued. As for the transformed wave functions, they change their sign when φ increases by 2π since the operators of the transformations, (2.11), are linear in $\sin(\varphi/2)$ and $\cos(\varphi/2)$. Hence the transformed functions are two-valued functions of position coordinates.

The eigenvalue equations (2.8) and (2.9) for $\mathcal{M}_z$ and $\mathcal{M}$ then read

$$p_\varphi \psi^* = \hbar \left(m + \frac{1}{2} \right) \psi^* \tag{2.28}$$

$$-\frac{\sigma_1}{\sin\theta} p_\varphi \psi^* + \sigma_2 p_\theta \psi^* = \hbar k \psi^* \tag{2.29}$$

3. Spherical harmonics with spin

If we write the two component wave function as

$$\psi^* = \begin{pmatrix} Y \\ Z \end{pmatrix} \tag{3.1}$$

we see that Eq. (2.28) splits into two similar equations

$$\frac{\partial Y}{\partial \varphi} = i \left(m + \frac{1}{2} \right) Y, \qquad \frac{\partial Z}{\partial \varphi} = i \left(m + \frac{1}{2} \right) Z \tag{3.2}$$

and Eq. (2.29) reduces to the system

$$\frac{i}{\sin\theta} \frac{\partial Z}{\partial \varphi} - \frac{\partial Z}{\partial \theta} = kY, \qquad \frac{i}{\sin\theta} \frac{\partial Y}{\partial \varphi} + \frac{\partial Y}{\partial \theta} = kZ \tag{3.3}$$

If we want the solutions to satisfy (3.2), we put

$$Y(\theta, \varphi) = \frac{1}{(4\pi)^{1/2}} e^{i(m+1/2)\varphi} A(\theta)$$

$$Z(\theta, \varphi) = \frac{1}{(4\pi)^{1/2}} e^{i(m+1/2)\varphi} B(\theta) \tag{3.4}$$

Then $A(\theta)$ and $B(\theta)$ will satisfy the following system of equations:

$$\frac{dA}{d\theta} = \frac{(m+1/2)}{\sin\theta} A + kB, \qquad \frac{dB}{d\theta} = -\frac{(m+1/2)}{\sin\theta} B - kA \tag{3.5}$$

and the normalization condition will be

$$\frac{1}{2} \int_0^\pi (A^2 + B^2) \, d\theta = 1 \tag{3.6}$$

These last functions can be expressed in terms of the associated Legendre polynomials of the first kind, $P_l^m(\cos\theta)$, studied in Schrödinger's theory (see Sections 4-6, Chapter IV, Part II). Let us recall some of their properties. The function

$$P_l^m(x) = P_l^m(\cos\theta) \tag{3.7}$$

satisfies the equation

$$\frac{d}{dx}\left[(1-x^2)\frac{dP_l^m}{dx}\right] - \frac{m^2}{1-x^2} P_l^m + l(l+1)P_l^m = 0 \tag{3.8}$$

and is the solution that remains finite at $x = \pm 1$. At $m = 0$ the function $P_l^m(x)$ reduces to the Legendre polynomial

$$P_l(x) = \frac{1}{2^l l!} \frac{d^l}{dx^l} (x^2 - 1)^l \tag{3.9}$$

and at $m \geqslant 0$ it is

$$P_l^m(x) = \left(1 - x^2\right)^{m/2} \frac{d^m}{dx^m} P_l(x) \tag{3.10}$$

For negative values of m the $P_l^m(x)$ are expressed in terms of the associated Legendre polynomials with positive values of m:

$$P_l^{-m}(x) = (-1)^m \frac{(l-m)!}{(l+m)!} P_l^m(x) \tag{3.11}$$

We also know [see (5.18) and (5.19), Chapter IV, Part II] that $P_l^m(\cos\theta)$ and $P_l^{m+1}(\cos\theta)$, considered as functions of θ, satisfy

the following system of first-order equations:

$$\frac{d}{d\theta} P_l^m (\cos\theta) - m \cot\theta\, P_l^m (\cos\theta) = - P_l^{m+1} (\cos\theta)$$

$$\frac{d}{d\theta} P_l^{m+1} (\cos\theta) + (m+1) \cot\theta P_l^{m+1} (\cos\theta)$$

$$= (l+m+1)(l-m) P_l^m (\cos\theta) \qquad (3.12)$$

Equations (3.5) reduce to this system. We can show this by initiating a transformation

$$A + iB = (\sin\theta)^{1/2} e^{-i\theta/2} (y_1 + iy_2) \qquad (3.13)$$

analogous to the one from ψ^* to ψ' given by (2.18) and (2.21) [we consider A, B, y_1, and y_2 in (3.13) to be real]. Then the equations for y_1 and y_2 are

$$\frac{dy_1}{d\theta} - m \cot\theta\, y_1 = (k+m) y_2$$

$$\frac{dy_2}{d\theta} + (m+1) \cot\theta\, y_2 = -(-k+m+1) y_1 \qquad (3.14)$$

These coincide with the equations for ordinary spherical harmonics, (3.12), if we put

$$y_1 = - c(k+m) P_l^m, \qquad y_2 = c P_l^{m+1} \qquad (3.15)$$

or

$$y_1 = - c' P_l^{-m}, \qquad y_2 = c'(-k+m+1) P_l^{-m-1} \qquad (3.16)$$

and if we assume, following (1.20), that

$$k(k-1) = l(l+1) \qquad (3.17)$$

We can take l to be either of the two numbers $-k$ and $k-1$ that is nonnegative. This condition can be written

$$l + \tfrac{1}{2} = |k - \tfrac{1}{2}| \qquad (3.18)$$

Equating (3.15) with (3.16) and using formula (3.11), we get the ratio of the constants c' and c:

$$\frac{c'}{c} = (-1)^m (k+m) \frac{(l+m)!}{(l-m)!} \qquad (3.19)$$

The value of each constant can be found from the normalization condition (3.6), which can be rewritten thus:

$$\frac{1}{2} \int_0^\pi (y_1^2 + y_2^2) \sin\theta\, d\theta = 1 \qquad (3.20)$$

We get

$$c = \frac{1}{|k+m|^{1/2}} \left(\frac{(l-m)!}{(l+m)!} \right)^{1/2}, \qquad c' = (-1)^m \frac{k+m}{|k+m|^{1/2}} \left(\frac{(l+m)!}{(l-m)!} \right)^{1/2}$$

(3.21)

If we introduce the normalized spherical harmonics by (6.7), Chapter IV, Part II,

$$P_l^{*m}(x) = (2l+1)^{1/2} \left(\frac{(l-m)!}{(l+m)!} \right)^{1/2} P_l^m(x), \qquad x = \cos\theta \quad (3.22)$$

where

$$\frac{1}{2} \int_{-1}^{+1} [P_l^{*m}(x)]^2 \, dx = \frac{1}{2} \int_0^\pi [P_l^{*m}(\cos\theta)]^2 \sin\theta \, d\theta = 1 \quad (3.23)$$

then

$$y_1 = - \frac{k+m}{[(k+m)(2k-1)]^{1/2}} P_l^{*m}, \qquad y_2 = \left(\frac{k-m-1}{2k-1} \right)^{1/2} P_l^{*m+1} \quad (3.24)$$

where the square roots must be taken to be positive. Bearing in mind the relationship between k and l (3.18), we write these formulas as

$$y_1 = - \left(\frac{k+m}{2k-1} \right)^{1/2} P_{k-1}^{*m}, \quad y_2 = \left(\frac{k-m-1}{2k-1} \right)^{1/2} P_{k-1}^{*m+1}, \quad k > 0 \quad (3.25)$$

$$y_1 = \left(\frac{k+m}{2k-1} \right)^{1/2} P_{-k}^{*m}, \quad y_2 = \left(\frac{k-m-1}{2k-1} \right)^{1/2} P_{-k}^{*m+1}, \quad k < 0 \quad (3.25^*)$$

These formulas demonstrate that the functions y with negative k can be expressed in terms of functions with positive k:

$$y_1(-k, m, \theta) = - \left(\frac{k-m}{k+m+1} \right)^{1/2} y_1(k+1, m, \theta)$$

$$y_2(-k, m, \theta) = \left(\frac{k+m+1}{k-m} \right)^{1/2} y_2(k+1, m, \theta) \quad (3.26)$$

We list some spherical harmonics with spin for various values of k:

$k = +1$	$(l = 0)$	
$m = -1$	$y_1 = 0$	$y_2 = 1$
$m = 0$	$y_1 = -1$	$y_2 = 0$

$$k = -1 \qquad (l = 1)$$

$m = -1$	$y_1 = -\sin\theta$	$y_2 = \cos\theta$
$m = 0$	$y_1 = \cos\theta$	$y_2 = \sin\theta$

$$k = +2 \qquad (l = 1)$$

$m = -2$	$y_1 = 0$	$y_2 = -\sqrt{\dfrac{3}{2}}\,\sin\theta$
$m = -1$	$y_1 = \dfrac{1}{\sqrt{2}}\sin\theta$	$y_2 = \sqrt{2}\,\cos\theta$
$m = 0$	$y_1 = -\sqrt{2}\,\cos\theta$	$y_2 = \dfrac{1}{\sqrt{2}}\sin\theta$
$m = 1$	$y_1 = -\sqrt{\dfrac{3}{2}}\,\sin\theta$	$y_2 = 0$

4. Some properties of spherical harmonics with spin

If we write Eqs. (3.5) as

$$\left[-i\sigma_2 \frac{\partial}{\partial\theta} - \frac{(m + 1/2)}{\sin\theta}\,\sigma_1 \right] \binom{A}{B} = k \binom{A}{B} \qquad (4.1)$$

then these can be interpreted as the eigenvalue equations for the hermitian operator

$$\mathscr{L} = -i\sigma_2 \frac{\partial}{\partial\theta} - \frac{(m + 1/2)}{\sin\theta}\,\sigma_1 \qquad (4.2)$$

with the eigenvalue being k. It then follows that functions A and B are orthogonal, that is

$$\int_0^\pi [A(k, \theta)\,A(k', \theta) + B(k, \theta)\,B(k', \theta)]\,d\theta = 0, \qquad k \neq k' \quad (4.3)$$

and constitute a complete set.

Now let us shift from A and B to y_1 and y_2 via (3.13). The latter then also constitute a complete orthogonal set. Recalling their normalization condition, (3.20), we can write

$$\frac{1}{2} \int_0^\pi [y_1(k, m, \theta)\,y_1(k', m, \theta) + y_2(k, m, \theta)\,y_2(k', m, \theta)]\sin\theta\,d\theta = \delta_{kk'}$$

$$(4.4)$$

16*

or simply

$$\frac{1}{2} \int_0^\pi y(k, m, \theta) y(k', m, \theta) \sin \theta \, d\theta = \delta_{kk'} \tag{4.5}$$

where y denotes the set of two functions, y_1 and y_2.

An arbitrary pair of functions $u_1(\theta)$ and $u_2(\theta)$ which we denote by one symbol $u(\theta)$, can be expanded (after certain conditions are met) in a $y(k, m, \theta)$-series:

$$u(\theta) = \sum_k c(k) y(k, m, \theta) \tag{4.6}$$

Actually, (4.6) is two expansions:

$$u_p(\theta) = \sum_k c(k) y_p(k, m, \theta), \qquad p = 1, 2 \tag{4.6*}$$

where the expansion coefficients, $c(k)$, are the same in both cases. These coefficients can be found via the integral

$$c(k) = \frac{1}{2} \int_0^\pi y(k, m, \theta) u(\theta) \sin \theta \, d\theta \tag{4.7}$$

which in explicit form is

$$c(k) = \frac{1}{2} \int_0^\pi [y_1, (k, m, \theta) u_1(\theta) + y_2(k, m, \theta) u_2(\theta)] \sin \theta \, d\theta \tag{4.7*}$$

With further applications in mind we put

$$u(\theta) = y(k_0, m, \theta) \cos \theta \tag{4.8}$$

To evaluate integrals of type (4.7*) we express $y(k, m, \theta)$ in terms of ordinary spherical harmonics via (3.25) and (3.25*) and then use the recursion relation (6.11), Chapter IV, Part II. In fact, only three expansion coefficients $c(k)$ are nonzero, which means that (4.6) contains no more than three terms. If we write k instead of k_0, this expansion will be

$$y(k, m, \theta) \cos \theta = - \frac{2m+1}{4k^2-1} y(-k, m, \theta)$$
$$+ \frac{[(k+m)(k-m-1)]^{1/2}}{|2k-1|} y(k-1, m, \theta)$$
$$+ \frac{[(k-m)(k+m+1)]^{1/2}}{|2k+1|} y(k+1, m, \theta) \tag{4.9}$$

[We can verify this by expressing $y(k, m, \theta)$ in terms of ordinary spherical harmonics.]

Now suppose that

$$u(\theta) = y(k_0, m-1, \theta) \sin \theta \qquad (4.8^*)$$

in contrast to (4.8). After a similar procedure we find the expansion for (4.8):

$$y(k, m-1, \theta) \sin \theta = 2 \frac{(k^2 - m^2)^{1/2}}{4k^2 - 1} y(-k, m, \theta)$$

$$- \frac{[(k-m-1)(k-m)]^{1/2}}{2k-1} y(k-1, m, \theta)$$

$$+ \frac{[(k+m+1)(k+m)]^{1/2}}{2k+1} y(k+1, m, \theta) \quad (4.10)$$

This can be verified by using the recursion relation (6.12), Chapter IV, Part II, of which (4.10) is a generalization.

We note that (4.9) and (4.10) hold not only for y_1 and y_2 but for functions A and B as well, since they are interconnected through a linear transformation. The coefficients of this transformation do not depend on either k or m.

The method by which we derived the recursion relations, using the completeness condition, is sufficiently general. It can be applied for ordinary spherical harmonics and generalized Laquerre polynomials, considered in Part II.

5. The Pauli wave equation

If a particle of mass m and electric charge $-e$ is in an electromagnetic field with vector potential A_x, A_y, A_z and scalar potential Φ, the *Lagrange function (Lagrangian)* of nonrelativistic classical mechanics is

$$\mathscr{L} = \frac{1}{2} m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - \frac{e}{c}(\dot{x}A_x + \dot{y}A_y + \dot{z}A_z) + e\Phi \quad (5.1)$$

The generalized "momenta" conjugate to coordinates x, y, z are

$$p_x = \frac{\partial \mathscr{L}}{\partial \dot{x}}, \qquad p_y = \frac{\partial \mathscr{L}}{\partial \dot{y}}, \qquad p_z = \frac{\partial \mathscr{L}}{\partial \dot{z}} \qquad (5.2)$$

These do not coincide with the components of momentum

$$P_x = m\dot{x}, \qquad P_y = m\dot{y}, \qquad P_z = m\dot{z} \qquad (5.3)$$

but are related to them via the formulas

$$p_x = P_x - \frac{e}{c}A_x, \qquad p_y = P_y - \frac{e}{c}A_y, \qquad p_z = P_z - \frac{e}{c}A_z \quad (5.4)$$

The particle's energy is

$$E = \dot{x}p_x + \dot{y}p_y + \dot{z}p_z - \mathscr{L} = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - e\Phi \quad (5.5)$$

Expressing (5.5) in terms of generalized momenta, we get the Hamiltonian function of classical mechanics

$$H = \frac{1}{2m}\left[\left(p_x + \frac{e}{c}A_x\right)^2 + \left(p_y + \frac{e}{c}A_y\right)^2 + \left(p_z + \frac{e}{c}A_z\right)^2\right] - e\Phi \quad (5.6)$$

In the absence of a magnetic field the vector potential can be taken to be zero, which transforms (5.6) into

$$H = \frac{1}{2m}\left(p_x^2 + p_y^2 + p_z^2\right) - e\Phi \quad (5.7)$$

or

$$H = \frac{1}{2m}\left(p_x^2 + p_y^2 + p_z^2\right) + U(x, y, z) \quad (5.8)$$

where

$$U = -e\Phi \quad (5.9)$$

is the particle's potential energy.

We have already seen that in Schrödinger's theory the Hamiltonian can be obtained from the classical Hamiltonian function by substituting for the generalized momenta p_x, p_y, p_z the operators

$$p_x = -i\hbar\frac{\partial}{\partial x}, \qquad p_y = -i\hbar\frac{\partial}{\partial y}, \qquad p_z = -i\hbar\frac{\partial}{\partial z} \quad (5.10)$$

What happens when we deal with spin? We can build the operator

$$P = \sigma_x p_x + \sigma_y p_y + \sigma_z p_z \quad (5.11)$$

and use it in building the Hamiltonian.

Next we find the *anticommutator* of the operator

$$\mathcal{M} = \sigma_x m_x + \sigma_y m_y + \sigma_z m_z + \hbar \quad (5.12)$$

[see (1.15)] and P. The two operators *anticommute*, that is, the anticommutator equals zero:

$$\mathcal{M}P + P\mathcal{M} = 0 \quad (5.13)$$

To prove this we use properties (1.6) of $\sigma_x, \sigma_y, \sigma_z$ and the obvious commutation relations

$$m_y p_z - p_z m_y = p_y m_z - m_z p_y = i\hbar p_x$$
$$m_z p_x - p_x m_z = p_z m_x - m_x p_z = i\hbar p_y$$
$$m_x p_y - p_y m_x = p_x m_y - m_y p_x = i\hbar p_z \quad (5.14)$$

Here we will not proceed with the calculations but will note that they simplify considerably if we use spherical coordinates rather than (Cartesian) rectangular coordinates. This will be done in Section 6.

If we assume, as is done in Schrödinger's theory, that an electron has only those degrees of freedom that are associated with the motion of a mass point in ordinary three-dimensional space, then substituting the operators into (5.8) unequivocally brings us to the already studied Schrödinger expression for the Hamiltonian. But if for an electron we introduce new degrees of freedom, namely, the ones connected with spin, then there are new possibilities of transforming from quantities of classical mechanics to quantum mechanical operators.

Using the properties of the Pauli matrices $\sigma_x, \sigma_y, \sigma_z$ and the commutativity of p_x, p_y, p_z, we come to the following expression for the Hamiltonian (5.8):

$$H = \frac{1}{2m} (\sigma_x p_x + \sigma_y p_y + \sigma_z p_z)^2 + U(x, y, z) \qquad (5.15)$$

Hence introducing operator P, (5.11), gives nothing new. It is different when there is a magnetic field. Then the classical Hamiltonian function has the form (5.6) and the generalized momenta do not coincide with the components of momentum but are connected via (5.4). We write these as

$$P_x = p_x + \frac{e}{c} A_x, \qquad P_y = p_y + \frac{e}{c} A_y, \qquad P_z = p_z + \frac{e}{c} A_z \qquad (5.16)$$

If we consider these quantities as operators, they are not commutative. The commutation relations for them are

$$\frac{i}{\hbar} (P_y P_z - P_z P_y) = \frac{e}{c} \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} \right) = \frac{e}{c} \mathscr{H}_x$$

$$\frac{i}{\hbar} (P_z P_x - P_x P_z) = \frac{e}{c} \left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} \right) = \frac{e}{c} \mathscr{H}_y$$

$$\frac{i}{\hbar} (P_x P_y - P_y P_x) = \frac{e}{c} \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) = \frac{e}{c} \mathscr{H}_z \qquad (5.17)$$

where $\mathscr{H}_x, \mathscr{H}_y, \mathscr{H}_z$ are the magnetic field components. Because of this with the magnetic field the transformation from operators p to operators P gives different results depending on whether we initiated the transformation in (5.8) or in (5.15). If we shift from p_x, p_y, p_z to P_x, P_y, P_z in (5.8), we return to (5.6), which we denote by H^0:

$$H^0 = \frac{1}{2m} \left[\left(p_x + \frac{e}{c} A_x \right)^2 + \left(p_y + \frac{e}{c} A_y \right)^2 + \left(p_z + \frac{e}{c} A_z \right)^2 \right] - e\Phi \qquad (5.18)$$

But if we do this in (5.15) and then use the relations (1.6) for matrices $\sigma_x, \sigma_y, \sigma_z$ and the commutation relations (5.17), we arrive

at the operator

$$H^* = H^0 + \mu^0 \left(\sigma_x \mathscr{H}_x + \sigma_y \mathscr{H}_y + \sigma_z \mathscr{H}_z \right) \tag{5.19}$$

For the sake of brevity we have put

$$\mu^0 \equiv \frac{e\hbar}{2mc} = 0.927 \times 10^{-20} \text{ erg G}^{-1} \tag{5.20}$$

This can be considered as the magnetic moment of the electron (the *Bohr magneton*).

Hamiltonian (5.19) is the generalization of the Hamiltonian in Schrödinger's theory when there is a magnetic field acting on the electron (we ignore relativistic corrections). We will call H^* the *Pauli operator*, and the wave equation

$$H^* \psi = i\hbar \frac{\partial \psi}{\partial t} \tag{5.21}$$

the *Pauli wave equation*.

6. Operator P in spherical and cylindrical coordinates and its relation to $\mathscr{M}$

Introduction of the spin dependent operator

$$P = \sigma_x P_x + \sigma_y P_y + \sigma_z P_z \tag{6.1}$$

is an important point in transferring from the Schrödinger equation to the Pauli equation. Let us study the interrelation of the Pauli operator with $\mathscr{M}$, the operator of the spin-orbit scalar of angular momentum. In Section 5 we established that these two operators anticommute [see (5.13)]. To study this interrelation more closely it is convenient to transform both operators to spherical coordinates. For $\mathscr{M}$ we did this in Section 2. As for P, we will proceed in two steps: first we transform P to cylindrical coordinates, and then we transform the result to spherical coordinates.

Since vector potential A_x, A_y, A_z is covariant,

$$P_x = p_x + \frac{e}{c} A_x, \qquad P_y = p_y + \frac{e}{c} A_y, \qquad P_z = p_z + \frac{e}{c} A_z \tag{6.2}$$

transform in the same way as p_x, p_y, p_z. Hence it is sufficient to transform only p_x, p_y, p_z (assuming, for the time being, that the vector potential is zero) and then include A_x, A_y, A_z in the final formulas.

Let us introduce the cylindrical coordinates, ρ and φ, in the following way:

$$x = \rho \cos \varphi, \qquad y = \rho \sin \varphi, \qquad z = z \tag{6.3}$$

The partial derivatives of ψ with respect to both new and old coordinates are related thus:

$$\frac{\partial \psi}{\partial x} = \cos \varphi \frac{d\psi}{\partial \rho} - \frac{\sin \varphi}{\rho} \frac{\partial \psi}{\partial \varphi}$$

$$\frac{\partial \psi}{\partial y} = \sin \varphi \frac{\partial \psi}{\partial \rho} + \frac{\cos \varphi}{\rho} \frac{\partial \psi}{\partial \varphi} \tag{6.4}$$

whence

$$p_x = \cos \varphi p_\rho - \frac{\sin \varphi}{\rho} p_\varphi$$

$$p_y = \sin \varphi p_\rho + \frac{\cos \varphi}{\rho} p_\varphi \tag{6.4*}$$

whereas p_z does not change at all.

Next we write P as

$$P = \sigma_1 p_x + \sigma_2 p_y + \sigma_3 p_z \tag{6.5}$$

with σ_1, σ_2, σ_3 being the matrices (1.12). Substitution of (6.4*) into (6.5) yields

$$P = (\sigma_1 \cos \varphi + \sigma_2 \sin \varphi)\, p_\rho$$
$$+ \frac{1}{\rho} (-\sigma_1 \sin \varphi + \sigma_2 \cos \varphi)\, p_\varphi + \sigma_2 p_z \tag{6.6}$$

The last expression can be considerably simplified if we choose an appropriate canonical transformation

$$\psi' = S\psi, \qquad P' = SPS^+ \tag{6.7}$$

We note that in Section 2 we used this transformation in studying the angular momentum operators. There we introduced the matrices

$$S = \cos \frac{\varphi}{2} + i\sigma_3 \sin \frac{\varphi}{2}, \qquad S^+ = \cos \frac{\varphi}{2} - i\sigma_3 \sin \frac{\varphi}{2} \tag{6.8}$$

With the help of these we can express the coefficients of p_ρ, p_φ, p_z in (6.6) in the following way:

$$\sigma_1 \cos \varphi + \sigma_2 \sin \varphi = S^+ \sigma_1 S$$

$$-\sigma_1 \sin \varphi + \sigma_2 \cos \varphi = S^+ \sigma_2 S$$

$$\sigma_3 = S^+ \sigma_3 S \tag{6.9}$$

[the last formulas are equivalent to (2.14)]. If we apply the transformation to operator (6.6) and put

$$p_\rho' = Sp_\rho S^+, \qquad p_\varphi' = Sp_\varphi S^+, \qquad p_z' = Sp_z S^+ \tag{6.10}$$

we can write

$$P' = \sigma_1 p_\rho' + \frac{1}{\rho} \sigma_2 p_\varphi' + \sigma_3 p_z' \tag{6.11}$$

Since matrix S does not contain coordinates ρ and z, it commutes with p_ρ and p_z, so that these operators do not change. As for operator p'_φ, we found it in Section 2 [formula (2.16)]. Applying this result, we get

$$p'_\rho = p_\rho, \qquad p'_\varphi = p_\varphi - \frac{\hbar}{2}\sigma_3, \qquad p'_z = p_z \qquad (6.12)$$

Hence operator P transformed to cylindrical coordinates is

$$P' = \sigma_1 p_\rho + \frac{1}{\rho}\sigma_2\left(p_\varphi - \frac{\hbar}{2}\sigma_3\right) + \sigma_3 p_z \qquad (6.13)$$

or, after we change $\sigma_2\sigma_3$ to $i\sigma_1$,

$$P' = \sigma_1\left(p_\rho - \frac{i\hbar}{2\rho}\right) + \frac{1}{\rho}\sigma_2 p_\varphi + \sigma_3 p_z \qquad (6.14)$$

To account for the vector potential we change p_ρ, p_φ, p_z in (6.14) to

$$P_\rho = p_\rho + \frac{e}{c}A_\rho, \qquad P_\varphi = p_\varphi + \frac{e}{c}A_\varphi, \qquad P_z = p_z + \frac{e}{c}A_z \quad (6.15)$$

where A_ρ, A_φ, A_z are the generalized components of the vector potential and satisfy the relationship

$$A_x\,dx + A_y\,dy + A_z\,dz = A_\rho\,d\rho + A_\varphi\,d\varphi + A_z\,dz \qquad (6.16)$$

We now find P in spherical coordinates. The transition from the cylindrical to the spherical coordinates

$$z = r\cos\theta, \qquad \rho = r\sin\theta \qquad (6.17)$$

is performed according to the formulas similar to those used in transferring from rectangular to cylindrical coordinates. By analogy with (6.4) and (6.4*) we get

$$\frac{\partial\psi}{\partial z} = \cos\theta\frac{\partial\psi}{\partial r} - \frac{\sin\theta}{r}\frac{\partial\psi}{\partial\theta}$$

$$\frac{\partial\psi}{\partial\rho} = \sin\theta\frac{\partial\psi}{\partial r} + \frac{\cos\theta}{r}\frac{\partial\psi}{\partial\theta} \qquad (6.18)$$

whence

$$p_z = \cos\theta\,p_r - \frac{\sin\theta}{2}p_\theta$$

$$p_\rho = \sin\theta\,p_r + \frac{\cos\theta}{2}p_\theta \qquad (6.18^*)$$

Substitution of these formulas for p_z and p_ρ into (6.14) yields the following expression for P':

$$P' = (\sigma_1\cos\theta - \sigma_3\sin\theta)\frac{1}{2}p_\theta + \frac{\sigma_2}{r\sin\theta}p_\varphi$$

$$+ (\sigma_1\sin\theta + \sigma_3\cos\theta)p_r - \frac{i\hbar}{2}\frac{\sigma_1}{r\sin\theta} \qquad (6.19)$$

Now, to simplify this we introduce a canonical transformation similar to (6.7):

$$\psi'' = T\psi', \qquad P'' = TP'T^+ \tag{6.20}$$

where matrices T and T^+ are

$$T = \cos\frac{\theta}{2} + i\sigma_2 \sin\frac{\theta}{2}, \qquad T^+ = \cos\frac{\theta}{2} - i\sigma_2 \sin\frac{\theta}{2} \tag{6.21}$$

The coefficients of the first three terms on the right of (6.19) can be expressed as

$$\sigma_1 \cos\theta - \sigma_3 \sin\theta = T^+\sigma_1 T$$

$$\sigma_2 = T^+\sigma_2 T$$

$$\sigma_1 \sin\theta + \sigma_3 \cos\theta = T^+\sigma_3 T \tag{6.22}$$

and σ_1 in the last term as

$$\sigma_1 = T^+ (\sigma_1 \cos\theta + \sigma_3 \sin\theta) T \tag{6.22*}$$

The new operator P'' then reads

$$P'' = \frac{\sigma_1}{r} p_\theta'' + \frac{\sigma_2}{r\sin\theta} p_\varphi'' + \sigma_3 p_r'' - \frac{i\hbar}{2r\sin\theta}(\sigma_1\cos\theta + \sigma_3\sin\theta) \tag{6.23}$$

where

$$p_\theta'' = Tp_\theta T^+, \qquad p_\varphi'' = Tp_\varphi T^+, \qquad p_r'' = Tp_r T^+ \tag{6.24}$$

Since T contains only one coordinate, θ, we have

$$p_\varphi'' = p_\varphi, \qquad p_r'' = p_r \tag{6.25}$$

whereas

$$p_\theta'' = p_\theta - i\hbar T \frac{\partial T^+}{\partial\theta} = p_\theta - \frac{\hbar}{2}\sigma_2 \tag{6.25*}$$

Substituting (6.25) and (6.25*) into (6.23), we get the final formula

$$P'' = \frac{\sigma_1}{r}\left(p_\theta - i\frac{\hbar}{2}\cot\theta\right) + \frac{\sigma_2}{r\sin\theta}p_\varphi + \sigma_3\left(p_r - \frac{i\hbar}{r}\right) \tag{6.26}$$

We note that operator P in both cylindrical and spherical coordinates acquires a somewhat simpler form if we effect the transformations

$$\psi^* = \rho^{1/2}\psi' \quad \text{for cylindrical coordinates} \tag{6.27}$$

$$\psi^* = r(\sin\theta)^{1/2}\psi'' \quad \text{for spherical coordinates} \tag{6.28}$$

Each of these corresponds to a specific transformation of P:

$$P^* = \rho^{1/2}P'\rho^{-1/2} \tag{6.29}$$

$$P^* = r(\sin\theta)^{1/2}P''r^{-1}(\sin\theta)^{-1/2} \tag{6.30}$$

After we make these transformations we get

$$P^* = \sigma_1 p_\rho + \frac{\sigma_2}{\rho} p_\varphi + \sigma_3 p_z \qquad (6.31)$$

for cylindrical coordinates, and

$$P^* = \frac{\sigma_1}{r} p_\theta + \frac{\sigma_2}{r \sin \theta} p_\varphi + \sigma_3 p_r \qquad (6.32)$$

for spherical coordinates.

We can also show that transformations (6.27) and (6.28) simplify the expressions for the probability of the electron's coordinates being within given limits. What we mean is that the probability for the inequalities

$$\rho < \rho' < \rho + d\rho, \quad \varphi < \varphi' < \varphi + d\varphi, \quad z < z' < z + dz \qquad (6.33)$$

is

$$|\psi'|^2 \rho \, d\rho \, d\varphi \, dz = |\psi^*|^2 \, d\rho \, d\varphi \, dz \qquad (6.34)$$

and the probability for the inequalities

$$\theta < \theta' < \theta + d\theta, \quad \varphi < \varphi' < \varphi + d\varphi, \quad r < r' < r + dr \qquad (6.35)$$

is (with a new value of ψ^*)

$$|\psi''|^2 r^2 \sin \theta \, d\theta \, d\varphi \, dr = |\psi^*|^2 \, d\theta \, d\varphi \, dr \qquad (6.36)$$

Now we are prepared to express P in terms of the operator of the spin-orbit scalar of angular momentum, $\mathscr{M}$. This is most easily done in spherical coordinates. Both operators were subjected to the same transformations: the transformation matrices S and T are the same for both operators [see (2.11) and (6.8) for S and (2.18) and (6.21) for T]. We can thereby deal with the transformed operators, P^* and $\mathscr{M}^*$:

$$P^* = \frac{\sigma_1}{r} p_\theta + \frac{\sigma_2}{r \sin \theta} p_\varphi + \sigma_3 p_r$$

$$\mathscr{M}^* = - \frac{\sigma_1}{\sin \theta} p_\varphi + \sigma_2 p_\theta \qquad (6.37)$$

We can easily see that

$$P^* = \sigma_3 \left(p_r + i \frac{\mathscr{M}^*}{r} \right) \qquad (6.38)$$

But the properties of matrices σ_1 and σ_2 imply that $\mathscr{M}^*$ anticommutes with σ_3. For this reason the equality

$$P^* = \left(p_r - i \frac{\mathscr{M}^*}{r} \right) \sigma_3 \qquad (6.39)$$

also holds. Here matrix σ_3 can be interpreted as the radial component of the spin. Indeed, if we put

$$\sigma_r = \frac{1}{r} (x \sigma_1 + y \sigma_2 + z \sigma_3) \qquad (6.40)$$

we will have

$$TS\sigma_r S^+ T^+ = \sigma_3 \tag{6.41}$$

so that after we have used in our formulas transformations S and T, matrix σ_r becomes σ_3. Hence prior to the application of S and T the formulas corresponding to (6.38) and (6.39) are

$$P = \sigma_r \left(p_r + i \frac{\mathcal{M}}{r} \right) = \left(p_r - i \frac{\mathcal{M}}{r} \right) \sigma_r \tag{6.42}$$

This can also be verified directly.

According to (5.13) operators P^* and $\mathcal{M}^*$ (as well as P and $\mathcal{M}$) are anticommutative. This is readily seen if we multiply (6.38) from the left by $\mathcal{M}^*$, then multiply (6.39) from the right by $\mathcal{M}^*$, and add the two products. We then get

$$\mathcal{M}^* P^* + P^* \mathcal{M}^* = p_r \left(\mathcal{M}^* \sigma_3 + \sigma_3 \mathcal{M}^* \right) = 0 \tag{6.43}$$

Finally, let us give the formula for P in curvilinear orthogonal coordinates. [1] If in coordinates q_1, q_2, q_3 the square of the element of arc length is

$$ds^2 = h_1^2 \, dq_1^2 + h_2^2 \, dq_2^2 + h_3^2 \, dq_3^2 \tag{6.44}$$

then

$$P = \frac{\sigma_1}{h_1} \left(p_1 - i \frac{\hbar}{2} \frac{\partial}{\partial q_1} (\log h_2 h_3) \right) + \frac{\sigma_2}{h_2} \left(p_2 - i \frac{\hbar}{2} \frac{\partial}{\partial q_2} (\log h_3 h_1) \right)$$
$$+ \frac{\sigma_3}{h_3} \left(p_3 - \frac{\hbar}{2} \frac{\partial}{\partial q_3} (\log h_1 h_2) \right) \tag{6.45}$$

where by p_k we have denoted the operators

$$p_k = - i\hbar \frac{\partial}{\partial q_k} \tag{6.46}$$

In the presence of an electromagnetic field we must change p_k to

$$P_k = p_k + \frac{e}{c} A_k \tag{6.47}$$

where the A_k are the covariant components of the vector potential and satisfy the condition

$$A_x \, dx + A_y \, dy + A_z \, dz = A_1 \, dq_1 + A_2 \, dq_2 + A_3 \, dq_3 \tag{6.48}$$

We readily see that in the particular cases of cylindrical and spherical coordinates formula (6.45) transforms into (6.14) and (6.26) respectively.

---

[1] This was done for a more general case in the author's work "The Dirac wave equation and Riemannian geometry" in *J. Russian Phys.-Chem. Sot.* 62: 133 (1930).

7. An electron in a magnetic field

Let us study the Pauli operator when there is a constant magnetic field. For the sake of clarity we will use Cartesian rectangular coordinates. If the magnetic field is sufficiently weak, in H^0 we can drop the terms that contain the square of the vector potential. In the terms that are linear in the vector potential we substitute for A_x, A_y, A_z the expressions

$$A_x = \frac{1}{2}(z\mathcal{H}_y - y\mathcal{H}_z)$$

$$A_y = \frac{1}{2}(x\mathcal{H}_z - z\mathcal{H}_x)$$

$$A_z = \frac{1}{2}(y\mathcal{H}_x - x\mathcal{H}_y) \tag{7.1}$$

This gives

$$A_x p_x + A_y p_y + A_z p_z = \frac{1}{2}(m_x\mathcal{H}_x + m_y\mathcal{H}_y + m_z\mathcal{H}_z) \tag{7.2}$$

where m_x, m_y, m_z are the components of the electron's orbital angular momentum [see (1.1)].

Using (7.2), we obtain the approximation for H^0:

$$H^0 = \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2) - e\Phi$$

$$+ \frac{e}{2mc}(m_x\mathcal{H}_x + m_y\mathcal{H}_y + m_z\mathcal{H}_z) \tag{7.3}$$

If to (7.3) we now add the spin dependent terms [see (5.19)], we get

$$H^* = \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2) - e\Phi$$

$$+ \frac{e}{2mc}[(m_x + \hbar\sigma_x)\mathcal{H}_x + (m_y + \hbar\sigma_y)\mathcal{H}_y + (m_z + \hbar\sigma_z)\mathcal{H}_z] \tag{7.4}$$

This expression contains the scalar product of the magnetic field by the vector of the electron's magnetic moment

$$\mu_x = \frac{e}{2mc}(m_x + \hbar\sigma_x)$$

$$\mu_y = \frac{e}{2mc}(m_y + \hbar\sigma_y)$$

$$\mu_z = \frac{e}{2mc}(m_z + \hbar\sigma_z) \tag{7.5}$$

This vector has two parts: the orbital and the spin part. The first is proportional to the electron's orbital angular momentum

$$m_x, \qquad m_y, \qquad m_z \tag{7.6}$$

whereas the second is proportional to the intrinsic angular momentum (spin)

$$\frac{1}{2}\,\hbar\sigma_x, \qquad \frac{1}{2}\,\hbar\sigma_y, \qquad \frac{1}{2}\,\hbar\sigma_z \qquad (7.7)$$

The ratio of the magnetic moment to the mechanical angular momentum of an electron, or the *gyromagnetic ratio*, associated with spin is twice the gyromagnetic ratio for orbital motion. This is sometimes called the "double magnetism" of electron spin.

In the central-field problem the correction term of H^* in (7.4), which depends on the magnetic field, commutes with the principal term, (5.7). Hence the correction to the energy level, due to the magnetic field, consists in adding an eigenvalue of the correction term in (7.4). If we choose the z axis directed along the field, the correction will be

$$\Delta E = \frac{e\hbar}{2mc}\,(m' \pm 1)\,\mathcal{H}_z \qquad (7.8)$$

where $\hbar m'$ is an eigenvalue of m_z.

The correction ΔE due to electron spin does not, however, bring any new levels into the picture since m' is an integer. Thus changing m' to $m' \pm 1$ results in nothing. It is the relativistic effects that here play an important role.

The Pauli operator, H^*, does not account for these corrections [see formula (7.4)]. If we include them, then even in a central field the radial equation contains not only the quantum number l of Schrödinger's theory but the quantum number k. The last is determined from the eigenvalue equation for spherical harmonics with spin.

$$\mathcal{M}\psi = k\hbar\psi \qquad (7.9)$$

[see (1.22)], and is connected with l in the following manner:

$$k\,(k-1) = l\,(l+1) \qquad (7.10)$$

[see (1.20)].

We know that at $l = 0$ there is only one value of k, and that is $k = 1$. But at $l = 1, 2, \ldots$ two values of k are possible, $k = l + 1$ and $k = -l$. As a result Schrödinger's energy level corresponding to a given value of l (and a given value of n, the principal quantum number) at $l \geq 1$ splits into two adjacent components, which form a doublet. This is customarily called the *relativistic doublet*.

In the radial equation the order of magnitude of the relativistic term, as compared with that of the principal term (the potential energy), is usually denoted by γ^2, where

$$\gamma = \frac{e^2}{\hbar c} = \frac{1}{137} \qquad (7.11)$$

is a dimensionless constant, called the fine-structure constant. In contrast, the influence of a magnetic field on an energy level is given by (7.8).

The splitting of energy levels in a magnetic field is called the *Zeeman effect.*

The exact theory of the Zeeman effect for the hydrogen atom will be dealt with at the end of this book on the basis of Dirac's theory of the electron. Here we only want to point out that the behaviour of an electron in a magnetic field proves convincingly that the electron has a new degree of freedom associated with spin.

The existence of this new degree of freedom plays an important role in the quantum theory of many-electron systems (atoms and molecules are examples). This theory cannot be formulated without accounting for the symmetry of the wave function with respect to all possible permutations of the electrons. The wave function, expressed in terms of the whole set of variables (x, y, z, σ) for each electron, must reverse its sign under interchange of any two sets of variables corresponding to two electrons. This condition is called the *Pauli exclusion principle*, or the *antisymmetric requirement* for the wave function. We must note here that the set of variables for each electron includes the spin variable, σ, in addition to the position variables. It is then evident that the need for spin arises already in the nonrelativistic theory.

The many-electron problem will be studied in Part IV.

THE MANY-ELECTRON PROBLEM
OF QUANTUM MECHANICS
AND THE STRUCTURE OF ATOMS

1. Symmetry properties of the wave function

In the previous chapters we studied wave functions for the states of one electron. If the state is stationary, the wave function must satisfy the Schrödinger equation. When we are dealing with states of a system of n electrons, the wave function must possess certain properties of symmetry (it must be antisymmetric) under interchange of the position and spin variables of the electrons. This requirement is called, as we know, the Pauli exclusion principle. More than that, in many cases the total spin (or the total intrinsic angular momentum) of the system of electrons is known. This results in an additional restriction on the wave function.

As we know, a one-electron wave function depends on the three coordinates x, y, z and the spin variable σ, which takes on two values only (for instance, $\sigma = +1$ and $\sigma = -1$). If we denote the set of the three coordinates by r, we can write a one-electron wave function as

$$\psi(x, y, z, \sigma) \equiv \psi(r, \sigma) \qquad (1.1)$$

Now we turn to the wave function of a system of electrons. The wave function depends on all the spatial and spin variables of the electrons. For n electrons we have

$$\psi = \psi(r_1, \sigma_1; r_2, \sigma_2; \ldots; r_n, \sigma_n) \qquad (1.2)$$

It is often convenient to denote by x_i all the variables referring to the ith electron (that is, the spatial and spin variables). Then the wave function for an n-electron system will be

$$\psi = \psi(x_1, x_2, \ldots, x_n) \qquad (1.3)$$

According to the Pauli exclusion principle, the wave function must be antisymmetric with respect to the variables $x_1, x_2, \ldots, x_n$, that is, it must change sign under interchange of any two variables. For instance

$$\psi(x_2, x_1, x_3, \ldots, x_n) = -\psi(x_1, x_2, x_3, \ldots, x_n) \qquad (1.4)$$

What are the requirements for a system of n electrons to have a definite value of total spin? First, we recall the main properties of spin studied in Part III.

When there is only one electron, any operator acting on the spin variable can be represented as a linear combination of the three operators $\sigma_x, \sigma_y, \sigma_z$. These are defined by the following relationships:

$$\sigma_x \psi (r,\ \sigma) = \psi (r,\ -\sigma)$$
$$\sigma_y \psi (r,\ \sigma) = -i\sigma \psi (r,\ -\sigma)$$
$$\sigma_z \psi (r,\ \sigma) = \sigma \psi (r,\ \sigma) \tag{1.5}$$

If we consider ψ to be a two component wave function, the first component being $\psi(r, +1)$ and the second $\psi(r, -1)$, the operators $\sigma_x, \sigma_y, \sigma_z$ act in the same way as the Pauli matrices

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{1.6}$$

The operators

$$s_x = \frac{1}{2}\,\sigma_x, \qquad s_y = \frac{1}{2}\,\sigma_y, \qquad s_z = \frac{1}{2}\,\sigma_z \tag{1.7}$$

satisfy the commutation relations

$$s_y s_z - s_z s_y = is_x$$
$$s_z s_x - s_x s_z = is_y$$
$$s_x s_y - s_y s_x = is_z \tag{1.8}$$

which are characteristic of angular momentum in general (in units of $\hbar$). Operators (1.7) can therefore be interpreted as the operators of the components of the electron's intrinsic angular momentum. For a many-electron system we can in the same way, (1.5), define operators $\sigma_{lx}, \sigma_{ly}, \sigma_{lz}$ that act on the spin variable of the lth electron. We have

$$\sigma_{lx}\psi = \psi (r_1\ \sigma_1; r_2,\ \sigma_2;\ \ldots;\ r_l,\ -\sigma_l;\ \ldots)$$
$$\sigma_{ly}\psi = -i\sigma_l\psi (r_1,\ \sigma_1; r_2,\ \sigma_2;\ \ldots;\ r_l,\ -\sigma_l;\ \ldots)$$
$$\sigma_{lz}\psi = \sigma_l\psi (r_1,\ \sigma_1; r_2,\ \sigma_2;\ \ldots;\ r_l,\ \sigma_l;\ \ldots) \tag{1.9}$$

The operators for the components of total spin angular momentum (in units of $\hbar$) can then be defined by analogy with (1.7) in the following way:

$$s_x = \frac{1}{2}\,(\sigma_{1x} + \sigma_{2x} + \ldots + \sigma_{nx})$$
$$s_y = \frac{1}{2}\,(\sigma_{1y} + \sigma_{2y} + \ldots + \sigma_{ny})$$
$$s_z = \frac{1}{2}\,(\sigma_{1z} + \sigma_{2z} + \ldots + \sigma_{nz}) \tag{1.10}$$

These satisfy the same commutation relations (1.8). We also see that commutation relations (1.8) allow for an operator

$$s^2 = s_x^2 + s_y^2 + s_z^2 \tag{1.11}$$

which is the square of total spin angular momentum. This last operator commutes with each of the operators s_x, s_y, s_z, and its eigenvalues are $s(s+1)$, with s being half of a nonnegative integer. If the number of electrons, n, is even, s is a positive integer or zero. But if n is odd, s takes on half-odd-integral values. In both cases the difference $n/2 - s \equiv k$ is a nonnegative integer; k can be interpreted as the number of electron pairs with compensated spin.

For a given value of s the eigenvalues of s_x, s_y, s_z run through a sequence of numbers

$$-s, \; -s+1, \; \ldots, \; s-1, \; s \tag{1.12}$$

that is, $2s + 1$ values in all.

We can represent operator s^2 in the following way:

$$s^2 = n - \frac{n^2}{4} + \sum_{i<j} P_{ij} \tag{1.13}$$

where by P_{ij} we have denoted a permutation of the spin variables σ_i and σ_j.

We can now formulate the condition for a system of n electrons to have a given ealue of total spin in the form of an equation:

$$s^2 \psi = s(s+1)\psi \tag{1.14}$$

What are the eigenfunctions of this equation? To determine these we put $k = n/2 - s$ and let α_1, α_2, $\ldots$, α_k be a set of k different numbers from the sequence 1, 2, $\ldots$, n. We introduce a function of the spin variables

$$F_{a_1 a_2 \ldots a_k} = F\left(\sigma_{a_1}\sigma_{a_2} \ldots \sigma_{a_k} \mid \sigma_{a_{k+1}}\sigma_{a_{k+2}} \ldots \sigma_{a_n}\right) \tag{1.15}$$

symmetric both in the arguments σ_{a_1}, σ_{a_2}, $\ldots$, σ_{a_k} that stand to the left of the vertical bar and in the arguments $\sigma_{a_{k+1}}$, $\sigma_{a_{k+2}}$, $\ldots$ $\ldots$, σ_{a_n} that stand to the right. We also introduce a set of $\binom{n}{k}$ functions of the position coordinates of all the electrons,

$$\psi_{a_1 a_2 \ldots a_k} = \psi(r_1, r_2, \ldots, r_n) \tag{1.16}$$

which do not include spin variables as arguments. Functions (1.16) and also functions (1.15) are symmetric in the labels

17*

$\alpha_1, \alpha_2, \ldots, \alpha_k$. We can now build the series

$$\psi(r_1, \sigma_1; r_2, \sigma_2; \ldots; r_n, \sigma_n)$$
$$= \sum_{(\alpha_1 \cdots \alpha_k)} \psi_{\alpha_1 \alpha_2 \ldots \alpha_k}(r_1, r_2, \ldots, r_n) F_{\alpha_1 \alpha_2 \ldots \alpha_k}(\sigma_1, \sigma_2, \ldots, \sigma_n)$$

$$(1.17)$$

On the basis of (1.13) we can show that the series (1.17) satisfies Eq. (1.14) with a value of s equal to $n/2 - k$, provided that the functions of position coordinates, (1.16), are restricted by the subsidiary conditions

$$\sum_\alpha \psi_{\alpha_1 \alpha_2 \ldots \alpha_k} = 0 \qquad (1.18)$$

where α runs through all values of the sequence $1, 2, \ldots, n$ except $\alpha_2, \ldots, \alpha_k$. The number of such conditions is $\binom{n}{k-1}$.

For the obtained eigenfunction of s^2 to be a possible physical state of the system of electrons with a given total spin we must see whether it satisfies the Pauli exclusion principle, that is, whether it is antisymmetric in x_i. This is so if we express all functions (1.16) via the formula

$$\psi_{\alpha_1 \alpha_2 \ldots \alpha_k}(r_1, r_2, \ldots r_n)$$
$$= \varepsilon(P)\psi\left(r_{\alpha_1}, r_{\alpha_2}, \ldots, r_{\alpha_k} \mid r_{\alpha_{k+1}}, r_{\alpha_{k+2}}, \ldots, r_{\alpha_n}\right) \quad (1.19)$$

in terms of one function of the electrons' position coordinates. namely

$$\psi = \psi(r_1, r_2, \ldots, r_k \mid r_{k+1}, r_{k+2}, \ldots, r_n) \qquad (1.20)$$

In (1.19) the labels $\alpha_1, \alpha_2, \ldots, \alpha_n$ are the numbers $1, 2, \ldots, n$ taken in an arbitrary order, and P is the permutation

$$P = \begin{pmatrix} 1 & 2 & \ldots & n \\ \alpha_1 & \alpha_2 & \ldots & \alpha_n \end{pmatrix} \qquad (1.21)$$

which replaces 1 by α_1, 2 by α_2, etc. By $\varepsilon(P)$ we will denote a number equal to $+1$ if permutation P is even and -1 if it is odd.

What condition must function (1.20) satisfy? It must

(1) be antisymmetric in the first k arguments [which in (1.20) stand to the left of the bar], that is, for instance

$$\psi(r_2, r_1, r_3, \ldots, r_k \mid r_{k+1}, \ldots, r_n)$$
$$= -\psi(r_1, r_2, r_3, \ldots, r_k \mid r_{k+1}, \ldots, r_n) \quad (1.22)$$

(2) be antisymmetric in the last $n - k$ arguments [to the right of the bar in (1.20)], that is, for instance

$$\psi(r_1, r_2, \ldots, r_k \mid r_{k+2}, r_{k+1}, r_{k+3}, \ldots, r_n)$$
$$= -\psi(r_1, r_2, \ldots, r_k \mid r_{k+1}, r_{k+2}, r_{k+3}, \ldots, r_n) \quad (1.23)$$

(3) possess the property of cyclic symmetry, which is expressed by the equality

$$\psi(r_1, \ldots, r_{k-1}, r_k \,|\, r_{k+1}, r_{k+2}, \ldots, r_n)$$

$$= \psi(r_1, \ldots, r_{k-1}, r_{k+1} \,|\, r_k, r_{k+2}, \ldots, r_n) +$$

$$\ldots + \psi(r_1, \ldots, r_{k-1}, r_{k+l} \,|\, r_{k+1}, \ldots, r_{k+l-1}, r_k, r_{k+l+1}, \ldots, r_n) +$$

$$\ldots + \psi(r_1, \ldots, r_{k-1}, r_n \,|\, r_{k+1}, \ldots, r_{n-1}, r_k) \qquad (1.24)$$

The right-hand side consists of $n - k$ terms. Each of these is obtained from the left-hand side by successively interchanging the argument r_k, lying to the left of the bar.

The cyclic symmetry is the result of the subsidiary conditions (1.18). Each of the $\binom{n}{k-1}$ conditions (1.18) corresponds to an equality of type (1.24). This can be verified by a direct computation, in which we must allow for the properties of antisymmetry (1.22) and (1.23). Suppose that by P we denote a cyclic permutation of the set of arguments $r_k, r_{k+1}, \ldots, r_n$, that is, a permutation in which each member of the set is replaced by the next member, the last member taking the position of the first. Then (1.24) can be rewritten as

$$(1 + P + P^2 + \ldots + P^{n-k})\psi = 0 \qquad (1.25)$$

for even values of $n - k$, and

$$(1 - P + P^2 - \ldots - P^{n-k})\psi = 0 \qquad (1.26)$$

for odd values.

In the particular case of two electrons the state with total spin zero ($n = 2$, $k = 1$) is described by a symmetric function of coordinates, and a state with total spin one ($n = 2$, $k = 0$) by an antisymmetric function.

A very important example of a function of n arguments $r_1, r_2, \ldots, r_n$ that satisfies the three symmetry conditions just formulated is the product of two determinants

$$\psi = \Psi^{(1)}\Psi^{(2)} \qquad (1.27)$$

where

$$\Psi^{(1)} = \begin{vmatrix} \psi_1(r_1) & \ldots & \psi_1(r_k) \\ \cdot & \cdot \cdot \cdot \cdot \cdot & \cdot \\ \psi_k(r_1) & \ldots & \psi_k(r_k) \end{vmatrix}$$

$$\Psi^{(2)} = \begin{vmatrix} \psi_1(r_{k+1}) & \ldots & \psi_1(r_n) \\ \cdot & \cdot \cdot \cdot \cdot \cdot \cdot & \cdot \\ \psi_{n-k}(r_{k+1}) & \ldots & \psi_{n-k}(r_n) \end{vmatrix} \qquad (1.28)$$

These determinants use the one-electron functions

$$\psi_1(r), \ \psi_2(r), \ \ldots, \ \psi_{n-k}(r) \qquad (1.29)$$

that depend on position coordinates only. The larger of the two determinants contains all of the $n - k$ functions (1.29), whereas the smaller contains only the first k functions.

With this reasoning we are able to express the wave function (1.2), which depends on the spin variables in addition to the position coordinates, in terms of the Schrödinger wave function (1.20), which depends only on the coordinates. In the process we strictly take into account the Pauli exclusion principle and the eigenvalue equation for spin angular momentum, (1.14).

Though the Schrödinger wave function does not depend on spin variables, its properties do depend on the value of total spin, since total spin affects the symmetry of the wave function.

This explains what at first glance seems a paradox, that neither the Schrödinger equation nor the wave function contains spin variables and yet the energy levels depend on the values of total spin. This paradox can be solved in the following way. There are additional symmetry conditions imposed on the wave function that corresponds to an energy level with given spin. And these conditions differ for different values of total spin.

2. The Hamiltonian and its symmetry

The wave function that describes a stationary state of a many-electron system must be an eigenfunction of the appropriate Hamiltonian. By analogy with classical mechanics the Hamiltonian can be written as

$$H = -\frac{\hbar^2}{2m} \sum_{k=1}^{n} \nabla_k^2 + \sum_{k=1}^{n} U(x_k, y_k, z_k)$$

$$+ \sum_{k>l=1}^{n} \frac{e^2}{[(x_k - x_l)^2 + (y_k - y_l)^2 + (z_k - z_l)^2]^{1/2}} \qquad (2.1)$$

Here ∇_k^2 is the Laplacian operator that acts on the kth electron; $U(x, y, z)$ is the potential energy of the external field, external in relation to the n electrons (for instance, the potential energy of the field produced by the nucleus of the atom or by the nuclei in the case of a molecule); and the double sum is the potential energy of interaction between the electrons. Hamiltonian (2.1) corresponds to the case without a magnetic field. If the system of electrons were in an external magnetic field, the Hamiltonian would contain terms that depend on spin.

The energy levels and the stationary states of the system are determined from the equation

$$H\psi = E\psi \qquad (2.2)$$

where H is given by (2.1). We have already noted before that though H does not contain spin variables, the energy levels E depend on the quantum number s (spin angular momentum). The explanation is that the symmetry properties of the Schrödinger wave function, ψ, depend on the values of s.

For an atom, H is spherically symmetric, that is, it does not change under rotations of the coordinate axes in space. We can then subject the Schrödinger wave function (a function of coordinates) to the condition that it simultaneously be the eigenfunction of the operator of the square of orbital angular momentum (the quantum number l) and that of the operator of one of the components of angular momentum (the quantum number m). If, in addition, this eigenfunction possesses symmetry properties that correspond to a definite value of s, then it can be used to build a function of type (1.2). This last is a wave function with spin s. It satisfies the Pauli exclusion principle and is a simultaneous eigenfunction of the following five operators: the Hamiltonian, the square of orbital angular momentum, the square of spin angular momentum, the square of total (spin and orbital) angular momentum, and the component of total angular momentum along a coordinate axis. The function is built by using the so-called vector model, but we will not elaborate on this.

The Hamiltonian of a diatomic molecule has not spherical but only axial symmetry (that is, it does not change under rotations about the axis that connects the two atoms). Axial symmetry too can be used to introduce quantum numbers and to partially determine the wave functions.

If we use the spherical or axial symmetry that a system possesses, we can introduce certain quantum numbers and thus classify the energy levels. But considerations of symmetry are not sufficient to determine the levels themselves and the stationary states. An exact solution of Eq. (2.2) presents insurmountable mathematical difficulties (with the exception of the case of one electron). For this reason the development of approximate methods is most important. The most fruitful of these is the *self-consistent field method*, which we will consider next.

3. The self-consistent field method

We can obtain the eigenvalue equation for the Hamiltonian by applying the variational method, that is, requiring that

$$\delta W = 0 \tag{3.1}$$

with

$$W = \frac{1}{N} \int \bar{\psi} H \psi \, dV, \qquad N = \int \bar{\psi} \psi \, dV \tag{3.2}$$

In these formulas we assume ψ to be the function of coordinates introduced in Section 1, that is, a function not dependent on spin variables. The volume element of configuration space, dV, is then the product of the differentials of all the coordinates of the electrons:

$$dV = dx_1 \, dy_1 \, dz_1 \, \ldots \, dx_n \, dy_n \, dz_n \qquad (3.3)$$

The normalization integral, N, can be considered a given constant.

To prove our statement we find the variations of the integrals in (3.2). Since H is hermitian,

$$\delta \int \bar{\psi} H \psi \, dV = \int \delta \bar{\psi} H \psi \, dV + \text{complex conjugate} \qquad (3.4)$$

Also

$$\delta \int \bar{\psi} \psi \, dV = \int \delta \bar{\psi} \psi \, dV + \text{complex conjugate} \qquad (3.5)$$

Then we multiply the second equality by a constant real quantity E, subtract the product from the first equality, and nullify the difference. We get

$$\int \delta \bar{\psi} \, (H\psi - E\psi) \, dV + \text{complex conjugate} = 0 \qquad (3.6)$$

This must hold for an arbitrary variation of the real and imaginary parts of ψ, which is only possible if the coefficient of $\delta \bar{\psi}$ in the integrand is zero. Whence

$$H\psi = E\psi \qquad (3.7)$$

which is the eigenvalue equation for the Hamiltonian. This ends the proof.

The physical meaning of W is that it is the mathematical expectation of the system's energy in state ψ. The extremal value of W is the energy level E. To obtain the lowest level with a given value of quantum number s we must, in varying the integral, choose for comparison only functions ψ that possess the "right" symmetry properties and satisfy some general conditions (the derivatives of ψ must exist and all integrals must have finite values). If we want to obtain the higher levels, we must in addition require that the wave function corresponding to these levels be orthogonal to the wave functions of the lower levels.

To simplify matters we can require that the wave function satisfy some additional conditions. For instance, we can choose the wave function in the form of a product of two determinants, which in turn consist of the one-electron wave functions [see (1.27) and (1.28)]. Of course, in this case we would obtain an energy level that lies somewhat higher than the true one. The difference, however, will be small. In the same manner we can obtain all the other energy levels (also approximately).

Let us calculate the result by substituting into W the product (1.28). We will assume that the functions $\psi_p(r)$ are orthogonal:

$$\int \overline{\psi_p(r)}\, \psi_q(r)\, d\tau = \delta_{pq}, \qquad d\tau = dx\, dy\, dz \tag{3.8}$$

which is true without loss of generality. For this we express the Hamiltonian (2.1) as

$$H(r_1, r_2, \ldots, r_n) = \sum_{p=1}^{n} H(r_p) + \sum_{p>q=1}^{n} \frac{e^2}{|r_p - r_q|} \tag{3.9}$$

where

$$H(r) = -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{dz^2}\right) + U(x, y, z) \tag{3.10}$$

Then

$$W = \sum_{p=1}^{k} \int \overline{\psi_p(r)}\, H(r)\, \psi_p(r)\, d\tau + \sum_{p=1}^{n-k} \int \overline{\psi_p(r)}\, H(r)\, \psi_p(r)\, d\tau$$

$$+ \frac{e^2}{2} \int\int \frac{\rho^{(1)}(r, r)\, \rho^{(1)}(r', r') - |\rho^{(1)}(r, r')|^2}{|\mathbf{r} - \mathbf{r}'|}\, d\tau\, d\tau'$$

$$+ \frac{e^2}{1} \int\int \frac{\rho^{(2)}(r, r)\, \rho^{(2)}(r', r') - |\rho^{(2)}(r, r')|^2}{|\mathbf{r} - \mathbf{r}'|}\, d\tau\, d\tau'$$

$$+ e^2 \int\int \frac{\rho^{(1)}(r, r)\, \rho^{(2)}(r', r')}{|\mathbf{r} - \mathbf{r}'|}\, d\tau\, d\tau' \tag{3.11}$$

where we have introduced the notations

$$\rho^{(1)}(r, r') = \sum_{p=1}^{k} \overline{\psi_p(r)}\, \psi_p(r') \tag{3.12}$$

$$\rho^{(2)}(r, r') = \sum_{p=1}^{n-k} \overline{\psi_p(r)}\, \psi_p(r') \tag{3.13}$$

The formulas obtained allow for a clear interpretation. First of all, we assume that the overall wave function can be expressed in terms of one-electron wave functions $\psi_p(r)$. This is equivalent to each electron having its own wave function (we can also say, its own *orbit*). The electrons in this case separate into two "clouds". The first corresponds to electrons on orbits ψ_1, ψ_2, ..., ψ_k, and the second to electrons on orbits ψ_1, ψ_2, ..., ψ_{n-k}. The two clouds have opposite spin angular momenta. Each orbit ψ_1, ψ_2, ... ψ_k "carries" two electrons with opposite spin angular momenta. The remaining orbits, ψ_{k+1}, ..., ψ_{n-k}, have one electron each, and the spin angular momenta of these electrons are the same. Thus for the first k orbits the total spin of each electron pair is zero. For the remaining $n - 2k$ orbits the spin angular momenta add up.

Since the electron spin (its absolute value) is $1/2$, the total spin of the system is $n/2 - k = s$, which is what we expected. We can interpret the e times $\rho^{(1)}(r, r)$, which is

$$e\rho^{(1)}(r, r) = e \sum_{p=1}^{k} |\psi_p(r)|^2 \tag{3.14}$$

as the *space charge density* of the first cloud [a similar interpretation can be given to $e\rho^{(2)}(r, r)$]. As for (3.12) and (3.13), they have different arguments r and r' and do not allow for a classical interpretation. We call them "mixed" charge densities.

Let us now interpret the expression for the energy of the electron system, (3.11). The first sum on the right is the kinetic and the potential energy of the first cloud in the field produced by the nuclei; the second sum is the same for the second cloud. The terms in first double integral that contain $\rho^{(1)}$ of the same arguments represent the electrostatic energy of the electrons of the first cloud. What remains in the first double integral contains the mixed charge density. It does not have a classical interpretation, and its presence in the energy is a specifically quantum effect (it is called the *exchange interaction*). The second double integral corresponds to the second electron cloud, and it can be interpreted in a similar manner. Finally, the third double integral corresponds to the electrostatic interaction between the two clouds.

Our interpretation is not rigorous. However, it is pictorial and hence useful for understanding the physical meaning of our formulas. In the strict sense formula (3.11) is the result of substituting a wave function having the "right" symmetry properties into the integral that is being varied.

If we vary (3.11) under conditions (3.8), we can find the system of equations for the sought functions $\psi_p(r)$. The system is

$$2[H(r) + V(r)]\psi_p(r) - e^2 \int \frac{[\rho^{(1)}(r', r) + \rho^{(2)}(r', r)]}{|\mathbf{r} - \mathbf{r}'|} \psi_p(r')\, d\tau'$$

$$= \sum_{q=1}^{n-k} \lambda_{qp}\psi_q(r), \qquad p = 1, 2, \ldots, k \tag{3.15}$$

$$[H(r) + V(r)]\psi_p(r) - e^2 \int \frac{\rho^2(r', r)}{|\mathbf{r} - \mathbf{r}'|} \psi_p(r')\, d\tau'$$

$$= \sum_{q=1}^{n-k} \lambda_{qp}\psi_q(r), \qquad p = k+1, \ldots, n-k \tag{3.16}$$

In both formulas

$$V(r) = e^2 \int \frac{\rho^{(1)}(r', r') + \rho^{(2)}(r', r')}{|\mathbf{r} - \mathbf{r}'|}\, d\tau' \tag{3.17}$$

can be interpreted as e times the potential of all the electrons. The quantities λ_{qp} are the Lagrange undetermined multipliers, which correspond to the orthogonality conditions (3.8). These conditions must be taken into account when building the variation δW. We can consider the off-diagonal elements of matrix λ_{qp} to be nonzero only if one of the labels is greater than or equal to $k + 1$ and the other is less than or equal to k.

We note that Eq. (3.16) for $\psi_p(r)$ does not contain the function in its coefficients, so that if we consider all the other functions to be known, the equation for $\psi_p(r)$ will be linear.

We can formulate this property of the equation in the following way. We put

$$\rho_p^{(2)}(r, r') = \sum_{q=1}^{n-k} (1 - \delta_{pq}) \overline{\psi_q(r)} \psi_q(r') \qquad (3.18)$$

and by $V_p(r)$ denote an expression similar to (3.17) but with sums (3.12) and (3.18) instead of (3.12) and (3.13). If we change V to V_p and $\rho^{(2)}$ to $\rho_p^{(2)}$, Eq. (3.16) retains its form.

Let us now turn to (3.15). We write this equation for the case when each orbit has two electrons, that is when $s = 0$. Then n is an even number, $k = n/2$, and the sums (3.12) and (3.13) coincide. Hence we can drop the upper index of ρ. More than that, we can assume in this case that λ_{qp} is a diagonal matrix and put

$$\lambda_{qp} = 2E_p \delta_{qp} \qquad (3.19)$$

As a result

$$[H(r) + V(r)] \psi_p(r) - e^2 \int \frac{\rho(r', r)}{|\mathbf{r} - \mathbf{r}'|} \psi_p(r') \, d\tau' = E_p \psi_p(r) \quad (3.20)$$

where

$$V(r) = 2e^2 \int \frac{\rho(r', r)}{|\mathbf{r} - \mathbf{r}'|} \, d\tau \qquad (3.21)$$

and

$$\rho(r', r) = \sum_{q=1}^{n/2} \overline{\psi_q(r')} \psi_q(r) \qquad (3.21^*)$$

If we then denote

$$V_p(r) = 2e^2 \int \frac{\rho(r', r')}{|\mathbf{r} - \mathbf{r}'|} \, d\tau' - e^2 \int \frac{|\psi_p(r')|^2}{|\mathbf{r} - \mathbf{r}'|} \, d\tau' \qquad (3.22)$$

and introduce a function similar to (3.18)

$$\rho_p(r', r) = \sum_{q=1}^{n-k} (1 - \delta_{pq}) \overline{\psi_q(r')} \psi_q(r)$$

we get Eq. (3.20) in the following form:

$$[H(r) + V_p(r)] \psi_p(r) - e^2 \int \frac{\rho_p(r', r)}{|\mathbf{r} - \mathbf{r}'|} \psi_p(r') \, d\tau' = E_p \psi_p(r)$$

Suppose that we drop the term with the integral. We then get the Schrödinger equation for an electron in a field with potential energy

$$\Phi = U(r) + V_p(r) \tag{3.23}$$

where $U(r)$ is the potential energy of the external field [see (3.10)], and $V_p(r)$ is the potential energy of the field produced by all the electrons except the given electron [$V_p(r)$ calculated by (3.22) is proportional to a potential that corresponds to charge density $\rho - |\psi_p|^2$]. As for the term with the integral, it has no classical counterpart. It is called the *correction for quantum exchange*.

Without the integral, Eq. (3.20) (a somewhat less accurate equation, in fact) was first suggested by the English mathematician D. R. Hartree. However, he did not give a satisfactory justification for such equations because in deriving them he used neither the variational method nor the idea of a wave function of the system as a whole. He proceeded from the obvious considerations just mentioned. Hartree called such equations *self-consistent field equations* (in the sense that the potential V in the equation for the wave functions is itself expressed in terms of these functions). Subsequently they were named the *Hartree equations*.

The complete equations with the integral terms, equations which take into account the symmetry of the system's wave function for a given value of spin angular momentum, have been obtained here by using the variational method. In doing so we have confirmed the Hartree equations. The complete equations have come to be known as *self-consistent field equations with quantum exchange*. In current literature these are also referred to as the *Hartree-Fock equations*.

The complete equations can be derived in another way, proposed by Dirac. This derivation differs in that the spin variables are included in the one-electron wave functions at the very outset (rather than excluded). The wave function (1.3) of a system of electrons is approximately expressed by one determinant of type

$$\Psi = \begin{vmatrix} \psi_1(x_1) & \dots & \psi_1(x_n) \\ \dots & \dots & \dots \\ \psi_n(x_1) & \dots & \psi_n(x_n) \end{vmatrix} \tag{3.24}$$

which contains the one-electron wave functions (1.1), called the *orbitals*. The resulting equations for these orbitals are similar to ours. What makes this method advantageous is the simple mathematics involved (since we have to do with one determinant instead of the product of two). The drawback is that Eq. (1.14) for the operator of spin angular momentum is not satisfied automatically but only after we have chosen the orbitals in a proper

way. When the problem possesses spherical symmetry, we can express the orbitals $\psi_i(x)$ in the determinant (3.24) in terms of the radial functions R_{nl} and the spherical harmonics with spin in such a way that the radial equations coincide with the ones we obtained by the previous method.

The self-consistent field equations for the orbitals can also be obtained from the theory of second quantization.

4. The equation for the valence electron and the operator of quantum exchange

We consider a system of an odd number of electrons, $n = 2k+1$, which has a spin equal to $^1/_2$ (for example, an atom with one valence electron). The overall wave function of such a system is a product of two determinants. It contains k wave functions ψ_1, ψ_2, ..., ψ_k, which enter into both determinants, and one wave function ψ, which enters into the larger determinant alone. We can say that functions ψ_1, ψ_2, ..., ψ_k describe $2k$ *inner-shell electrons* with compensated spin (two on each orbit) and function ψ describes the valence electron.

Using formula (3.11), we can write the energy of this system in the form of a sum:

$$W = W_0 + W' \qquad (4.1)$$

where

$$W_0 = 2 \sum_{p=1}^{k} \overline{\psi_p(r)} \, H(r) \, \psi_p(r) \, d\tau$$

$$+ e^2 \int \int \frac{2\rho(r, r)\,\rho(r', r') - |\rho(r, r')|^2}{|r - r'|} \, d\tau \, d\tau' \qquad (4.2)$$

is the energy of the inner-shell electrons, and

$$W' = \int \overline{\psi(r)} \, H(r) \, \psi(r) \, d\tau$$

$$+ e^2 \int \int \frac{2\rho(r', r')\,|\psi(r)|^2 - \rho(r', r)\,\overline{\psi(r)}\,\psi(r')}{|r - r'|} \, d\tau \, d\tau' \qquad (4.3)$$

is the energy of the valence electron of the field of the inner-shell electrons. The mixed charged density $\rho(r, r')$ is

$$\rho(r, r') = \sum_{p=1}^{k} \overline{\psi_p(r)} \, \psi_p(r') \qquad (4.4)$$

If we vary W over all the wave functions simultaneously, we will again arrive at the self-consistent field equations (3.15) and (3.16).

We can modify our problem and first determine the wave functions of the inner-shell electrons from the minimum condition

for W_0. Then, assuming that ψ_1, ψ_2, ..., ψ_k are given, we can determine the wave function of the outer electron, $\psi(r)$, from the minimum condition for W or W' (these two differ by the constant W_0).

This modified problem amounts, from the physical point of view, to first determining the stationary state of a system that contains one electron less (the *atomic core*) and then finding the state of the valence electron, ignoring the polarization of the atomic core by this electron.

For the wave functions of the inner-shell electrons we get a system of equations (3.20), and for the valence electron a linear integro-differential equation

$$[H(r) + V(r)] \psi(r) - e^2 \int \frac{\rho(r', r) \psi(r')}{|\mathbf{r} - \mathbf{r}'|} d\tau' = E \psi(r) \qquad (4.5)$$

where

$$V(r) = 2e^2 \int \frac{\rho(r', r')}{|\mathbf{r} - \mathbf{r}'|} d\tau' \qquad (4.6)$$

The wave function of the valence electron is always assumed to be orthogonal to the wave functions of the inner-shell electrons:

$$\int \overline{\psi_p(r)} \psi(r) \, d\tau = 0, \qquad p = 1, 2, \ldots, k \qquad (4.7)$$

It is easy to see that $\rho(r, r')$ and $V(r)$, determined in (4.4) and (4.6), coincide with (3.21*) and (3.21), and Eq. (4.5) with Eq. (3.20). Hence all the wave functions of the inner-shell electrons satisfy the equation for the valence electron as well; all are the eigenfunctions of the same linear integro-differential operator that stands on the left side of (4.5). It also follows from this that the orthogonality conditions (4.7) are automatically met (that is, these conditions are a corollary of the equation). As for the parameters E_p in (3.20), these are the eigenvalues of the same operator and can be interpreted as the energy levels of the inner-shell electrons.

Let us introduce a linear integro-differential operator $\mathscr{A}$, defining it as

$$\mathscr{A}\psi(r) = e^2 \int \frac{\rho(r', r)}{|\mathbf{r} - \mathbf{r}'|} \psi(r') \, d\tau' \qquad (4.8)$$

Recalling formula (3.10) for the Hamiltonian, $H(r)$, we can write our basic equation (4.5) as

$$-\frac{\hbar^2}{2m} \nabla^2 \psi + [U(r) + V(r)] \psi - \mathscr{A}\psi = E\psi \qquad (4.9)$$

The operator $-\mathscr{A}$ enters as a separate term into the expression for the energy of the electron. It can therefore be interpreted as a special kind of energy, usually called the *quantum exchange*

energy. If we examine the way that we derived Eq. (3.13), we see that the term $\mathscr{A}\psi$ appears because we took into account the symmetry properties of the wave function. These properties as well as the Pauli exclusion principle are connected with the indistinguishability of electrons and the impossibility of following any electron when it interacts with other electrons (the impossibility of labeling an electron so as to spot it after the interaction). This is an impossibility characteristic of quantum mechanics but not of classical mechanics (where the concept of a path is considered universally applicable). For this reason it is difficult to find a pictorial interpretation and a name for $\mathscr{A}$. The accepted term "quantum exchange" comes from the idea that when electrons interact, they change places, so to say.

5. The self-consistent field method in the theory of atoms

The simplest of the many-electron systems is the atom. What makes it easier to apply the self-consistent field method in describing various atoms is that even before quantum mechanics was fully established as we know it today (Schrödinger's theory) Bohr depicted the structure of electron shells for all atoms in the Periodic Table. His approach makes it possible to assign certain quantum numbers to each electron in the atom. These numbers are similar to those that characterize the state of a separate electron in a central field in the one-body problem. Bohr based his approach on experimental data, namely, the analysis of the spectra and chemical properties of atoms. Bohr's picture was confirmed theoretically by Schrödinger's quantum mechanics and the development of approximate methods, notably the self-consistent field method.

In quantum mechanics the possibility of ascribing specific quantum numbers to each electron in the atom implies the possibility of ascribing a specific wave function to each one. This is the very assumption that lies at the basis of the self-consistent field method, where the overall wave function is expressed in terms of one-electron wave functions. Thus we can determine the general character of the wave function if we know the quantum numbers of a given electron.

According to Bohr's approach the electrons in the atom separate into groups of *equivalent electrons.* Each such group is characterized by two quantum numbers: the principal quantum number n and the azimuthal quantum number l, where $n = 1, 2, 3, \ldots$ and $l = 0, 1, \ldots, n - 1$. Inside each group the electrons can be characterized by two other quantum numbers: the magnetic quantum number m and the spin quantum number m_s, where

$m = -l, -l + 1, \ldots, l - 1, l$ and $m_s = \pm^1/_2$. In the theory elaborated in Section 3 we took into account two values of the spin quantum number by separating the electrons into two clouds. Hence to characterize the wave function of an electron with given n and l it is sufficient to specify the value of the magnetic quantum number m.

If we denote the spherical coordinates with the origin in the nucleus by r, θ, φ, then the wave function of an electron with quantum numbers n, l, m will be

$$\psi_{nlm} = \frac{1}{\sqrt{4\pi}} R_{nl}(r) Y_{lm}(\theta, \varphi) \qquad (5.1)$$

where Y_{lm} is a normalized spherical harmonic:

$$\int |Y_{lm}(\theta, \varphi)|^2 \sin \theta \, d\theta \, d\varphi = 4\pi \qquad (5.2)$$

We must note that the choice of the quantum numbers that serve to separate the electrons inside a group is somewhat arbitrary because we can choose any direction as the polar axis. This arbitrariness is not evident, however, when a group of equivalent electrons is completely *filled*, that is, when the electrons in the group have all possible values of m. We say then that we have a *closed electron shell*.

Let us write the expression for the mixed charged density of the electrons of one cloud that belong to a closed electron shell. By (4.4) we have

$$\rho_{nl}(r, r') = \sum_{m=-l}^{+l} \overline{\psi_{nlm}(r, \theta, \varphi)} \, \psi_{nlm}(r', \theta', \varphi') \qquad (5.3)$$

Substituting ψ_{nlm} from (5.1) and using the addition theorem for spherical harmonics, we get

$$\rho_{nl}(r, r') = \frac{2l + 1}{4\pi} R_{nl}(r) R_{nl}(r') P_l(\cos \gamma) \qquad (5.4)$$

where

$$\cos \gamma = \cos \theta \cos \theta' + \sin \theta \sin \theta' \cos (\varphi - \varphi') \qquad (5.5)$$

and P_l is a Legendre polynomial. Since γ is the angle between the directions (θ, φ) and (θ', φ'), it does not depend on the choice of the polar axis. Formula (5.5) prompts the conclusion that a closed electron shell is spherically symmetric.

The number of electrons of one cloud that belong to a closed shell as $2l + 1$. Hence the total number of electrons of such a shell is $4l + 2$.

According to Bohr's picture most of the electrons in an atom make up closed shells. The rest, called the *outer electrons*, are situated in *unfilled shells*. All elements have only one unfilled

shell, with the exception of the rare earths and a few other elements, which have two.

The simplest case is when there is only one (valence) electron outside the closed shells. An example of such an atom is the sodium atom.

When studying problems connected with the atom, it is convenient to use a "natural" system of units (atomic units) introduced by Hartree [see Section 2, Chapter V, Part II]. In this system the electron charge, the electron mass, and Planck's constant h divided by 2π are put equal to unity. The atomic unit of length will be equal to 0.529 angstrom, the atomic unit of speed is the 137th part of the speed of light, and the atomic unit of energy will be twice the ground-state energy of the hydrogen atom, that is, 27.21 eV (1 hartree).

The Hamiltonian for sodium in atomic units is

$$H(r) = -\frac{1}{2} \nabla^2 - \frac{11}{r} \qquad (5.6)$$

since the charge of the nucleus of sodium is equal to 11 units. The inner-shell electrons of a sodium atom are grouped in three closed shells. We denote the corresponding wave functions as

$$\psi_1 = \psi_{100}, \quad \psi_2 = \psi_{200}, \quad \psi_3 = \psi_{21-1}, \quad \psi_4 = \psi_{210}, \quad \psi_5 = \psi_{211} \qquad (5.7)$$

where the ψ_{nlm} are given by (5.1). The mixed charge density of all three shells will be

$$\rho(r, r') = \frac{1}{4\pi} [R_{10}(r) R_{10}(r')$$
$$+ R_{20}(r) R_{20}(r') + 3R_{21}(r) R_{21}(r') \cos \gamma] \qquad (5.8)$$

Substituting this into the expression for W_0, (4.2), we get the energy of an ionized sodium atom. In the expression obtained for W_0 we can integrate over all angles. The expression will then depend only on the radial functions R_{10}, R_{20}, and R_{21}.

Now we can proceed in two ways, numerically or analytically. To apply the numerical method we must build equations involving variations of the radial functions. These equations have a form similar to (3.20). The numerical solution of *variational equations* is based on the method of successive approximations. We can then compile tables of radial functions with any desired degree of precision. To apply this method successfully it is important to choose a "reasonable" initial *trial* function. This is best done analytically. For this we look for R_{10}, R_{20}, and R_{21} in the form of analytic functions that depend on a small number of *variational parameters*. For instance, we can put

$$R_{10} = ae^{-\alpha r}, \qquad R_{20} = b\left(1 - \frac{1}{3}(\alpha + \beta) r\right), \qquad R_{21} = cre^{-\gamma r} \qquad (5.9)$$

where *a, b,* and *c* are found from the normalization condition

$$\int_0^\infty [R_{nl}(r)]^2 r^2\, dr = 1 \qquad (5.10)$$

and are expressed in terms of $\alpha, \beta,$ and γ. If we substitute (5.9) into the expression for W_0, we find W_0 as a linear-fractional function of α, β, and γ. By equating with zero the derivatives of W_0 with respect to parameters α, β, γ we arrive at an equation for α, β, γ. After solving these we get

$$\alpha = 10.68, \qquad \beta = 4.22, \qquad \gamma = 3.49 \qquad (5.11)$$

with the following radial functions (in analytic form):

$$R_{10}(r) = 69.804 e^{-10.68r}$$
$$R_{20}(r) = 13.602\,(1 - 4.967r)\,e^{-4.22r}$$
$$R_{21}(r) = 26.276 r e^{-3.49r} \qquad (5.12)$$

If we take these as the initial trial functions, we can get still better approximations from the variational equations. Functions (5.12) correspond to a definite value of the energy of an ionized sodium atom, namely, to $W_0 = -160.9$ hartrees, whereas the more exact functions, found by numerical solution of the equations, yield $W_0 = -161.8$ hartrees.

By finding the functions R_{10}, R_{20}, and R_{21} we have, in the given approximation, fully determined the electron-shell structure for sodium. We can now build the integro-differential equation (4.5) for the wave function of the valence electron. This function will be of type (5.1), where $R_{nl}(r)$ is found by solving numerically a certain equation, which we will not elaborate on here.

Solution of Eq. (4.5) first of all gives the optical terms of sodium. It also gives a rough estimate of the X-ray terms. To exhibit the accuracy of our method we offer the following table:

	E_{10}	E_{20}	E_{21}	E_{30}	E_{31}
Theory	−40.6	−3.00	−1.83	−0.1860	−0.1094
Experiment	−39.4	−2.16	−1.04	−0.1885	−0.1115

All figures are given in hartrees. We see that the optical terms (E_{30} and E_{31}) are found with a fair degree of accuracy. For instance, the error in E_{31} amounts to only 1.9 precent. For comparison, if in Eq. (4.5) we neglect the term with the integral, the

value of E_{31} is —0.08895, which corresponds to an error of 22.4 percent.

Apart from the optical terms our integro-differential equation makes it possible to find the wave functions for the valence electron on different energy levels. This enables us in turn to find the transition probabilities. For sodium and lithium the result agrees qualitatively with experiments. The result for lithium is the characteristic nonmonotonic dependence of the probability of the transition $E_{nl} \rightarrow E_{20}$ on quantum number n.

These results show that we can formulate the theory of the valence electron in an atom, with a high degree of accuracy, as the problem of a single body in a given field provided that we take into account the quantum exchange by introducing the integro-differential equation (4.5).

One of the ways in which we can use our equation has to do with what is known as the *sum rule*. This rule refers to the *oscillator strengths*, which are quantities proportional (a) to the square of the matrix element of the given transition and (b) to the difference in the two energy levels involved in this transition. The rule is usually stated in this way:

The sum of the oscillator strengths corresponding to all optical transitions in a given series must be unity.

This is an exact rule for atoms with one electron (*hydrogen-like atoms*, or ions with no electrons except one). As for atoms with a single valence electron, experiment has shown that in some cases (lithium, thallium and other atoms) the very first "oscillators" yield a sum greater than unity. Our theory explains this by noting that Eq. (4.5) has a complete set of eigenfunctions, and some of these functions correspond to X-ray terms. Hence, when building the sum of the oscillator strengths, we must account for the fictitious oscillators that correspond to transitions to occupied X-ray terms, which lie below the optical terms.

Since for these oscillators the difference in the energy levels that enters into the oscillator strength is negative, the oscillator strength itself will be negative. Thus in the expression for the total sum there will be a (finite) number of negative terms. It is clear from this that if the total sum is unity, the sum of the positive terms, which corresponds to the observed optical transitions, will exceed unity. We must also bear in mind that the total sum is unity only if we neglect certain small corrections due to quantum exchange.

Our equations in relation to atoms with a single valence electron are mainly applied in calculating energy levels and transition intensities. But attempts have also been made to use these

equations to account for some relativistic effects, namely, the distance between the two terms of a doublet (doublet splitting). Because of the difficulties involved in determining the wave function of the valence electron over small distances from the nucleus these attempts failed to give good quantitative results. For the doublet splitting of the spectral terms of sodium and lithium they only made it possible to estimate the order of magnitude. (We note that earlier calculations involving rougher approximations of the wave function did not even give the correct order of magnitude.) But they did bring out how important a role is played in the formula for doublet splitting by the terms responsible for the quantum exchange. It well may be that these terms explain the observed negative value of the doublet splitting for some atoms. In all likelihood, however, it will not be sufficient to consider the valence electron as being in the static field of the inner-shell electrons if we are to build an exact theory of doublet splitting.

6. The symmetry of the Hamiltonian of a hydrogen-like atom

In the previous section, following Bohr, we characterized each atomic shell by two quantum numbers, n and l. For a given value of n the number l takes on the values $l = 0$, 1, 2, ..., $n - 1$, altogether n values. All n electron shells belonging to a given value of n form the so-called "big" shell. (In fact, we often speak of electron *shells* and *subshells*. There is one shell for each principal quantum number n, and there is one subshell defined for each value of l allowed by $l = 0$, 1, 2, ..., $n - 1$. Thus each shell consists of a number of subshells.) This big shell possesses great stability in the atom. In the monovalence atoms of lithium, sodium, and copper there are one, two, and three big shells that are completely filled (closed). It is most convenient to describe a big shell with the help of *hydrogenic wave functions* corresponding to a certain effective charge of the nucleus. This charge can be determined using the variational method.

Let us denote the true charge of the nucleus by Z and the *effective charge* for a big electron shell with quantum number n by Z_n. It is more convenient to deal with the quantity $p_n = Z_n/n$, which is the *root-mean-square of the momentum* of an electron in the nth big shell (in atomic units), rather than with the effective charge Z_n.

Describing atomic shells with the help of (analytic) hydrogenic functions makes it possible to find simple formulas for the various functions that characterize the properties of an atom. For instance, the momentum distribution function for the electrons in the nth

shell, normalized according to the condition

$$4\pi \int_0^\infty \rho_n(\mathbf{p})\, p^2\, d\mathbf{p} = n^2 \tag{6.1}$$

appears to be

$$\rho_n(\mathbf{p}) = \frac{8p_n^3 n^2}{\pi^2 \left(p_n^2 + \mathbf{p}^2\right)^4} \tag{6.2}$$

The theory of hydrogenic functions can be applied not only to the theory of the electron-shell structure of atoms but also to the Compton scattering on bound electrons and other similar problems dealing with wave functions belonging to a continuous spectrum. For the hydrogen atom we have, in atomic units,

$$n = (-2E)^{-1/2} \tag{6.3}$$

In the continuous spectrum E is positive and (6.3) is pure imaginary. This, however, does not exclude the possibility of using the theory of hydrogenic functions belonging to a discrete spectrum for the case when the relationships obtained with their help are formulated in terms of analytic functions of n. We can then make a formal transition in these relationships from the discrete spectrum to the continuous, giving n pure imaginary values. For this reason we will confine ourselves to a discrete spectrum.

To write the Schrödinger equation in momentum space for a hydrogen-like atom we must first determine what corresponds to the operator of multiplication into $1/r$ in momentum space. The wave functions in coordinate space and momentum space are connected by the relation

$$\psi(p_x,\ p_y,\ p_z) = \frac{1}{(2\pi\hbar)^{3/2}} \int e^{-i(xp_x+yp_y+zp_z)/\hbar} \varphi(x,\ y,\ z)\, dx\, dy\, dz \tag{6.4}$$

We must find the form of the operator L that transforms function ψ into function $L\psi$, which can be represented as

$$L\psi(p_x,\ p_y,\ p_z) = \frac{1}{(2\pi\hbar)^{3/2}} \int e^{-i(xp_x+yp_y+zp_z)/\hbar} \frac{\varphi(x,\ y,\ z)}{r}\, dx\, dy\, dz \tag{6.5}$$

But

$$\frac{1}{2\pi^2\hbar} \int e^{-i(xp_x'+yp_y'+zp_z')/\hbar} \frac{d\mathbf{p}'}{|\mathbf{p}-\mathbf{p}'|^2} = \frac{1}{r} e^{-i(xp_x+yp_y+zp_z)/\hbar} \tag{6.6}$$

where $d\mathbf{p}' = dp_x'\, dp_y'\, dp_z'$ is the volume element in momentum space. Hence the operator that in coordinate space appears as multiplication into $1/r$ transforms in momentum space into the

integral operator

$$L\psi(p_x, p_y, p_z) = \frac{1}{2\pi^2\hbar} \int \frac{\psi(\mathbf{p}')}{|\mathbf{p} - \mathbf{p}'|^2} d\mathbf{p}' \qquad (6.7)$$

Thus the Schrödinger equation in a Coulomb field with potential energy $-Ze^2/r$ becomes in momentum space an integral equation of type

$$\frac{1}{2m} p^2 \psi(\mathbf{p}) = -\frac{Ze^2}{2\pi^2\hbar} \int \frac{\psi(\mathbf{p}')}{|\mathbf{p} - \mathbf{p}'|^2} d\mathbf{p}' \qquad (6.8)$$

Since we are dealing with the discrete spectrum, for which energy E is negative, we can introduce the root-mean-square momentum

$$p_0 = (-2mE)^{1/2} \qquad (6.9)$$

We will consider the components of momentum divided by p_0 as the rectangular coordinates on a hyperplane corresponding to the stereographic projection of a sphere of unit radius in four-dimensional Euclidean space. The rectangular coordinates of a point on the sphere are

$$\xi = \frac{2p_0 p_x}{p_0^2 + \mathbf{p}^2} = \sin\alpha \sin\theta \cos\varphi$$

$$\eta = \frac{2p_0 p_y}{p_0^2 + \mathbf{p}^2} = \sin\alpha \sin\theta \sin\varphi$$

$$\zeta = \frac{2p_0 p_z}{p_0^2 + \mathbf{p}^2} = \sin\alpha \cos\theta$$

$$\chi = \frac{p_0^2 - \mathbf{p}^2}{p_0^2 + \mathbf{p}^2} = \cos\alpha \qquad (6.10)$$

where

$$\xi^2 + \eta^2 + \zeta^2 + \chi^2 = 1 \qquad (6.11)$$

The angles α, θ, and φ are the spherical coordinates on the hypersphere. At the same time angles θ and φ are the angles that characterize the direction of momentum in three-dimensional space. The element of area on the hypersphere is

$$d\Omega = \sin^2\alpha \, d\alpha \sin\theta \, d\theta \, d\varphi \qquad (6.12)$$

and the total surface area of the hypersphere is $2\pi^2$. Instead of $\psi(\mathbf{p})$ we introduce the function

$$\Psi(\alpha, \theta, \varphi) = \frac{\pi}{\sqrt{8}} p_0^{-5/2} (p_0^2 + \mathbf{p}^2)^2 \psi(\mathbf{p}) \qquad (6.13)$$

for which the normalization condition is

$$\frac{1}{2\pi^2} \int | \Psi (\alpha, \theta, \varphi) |^2 \, d\Omega = \int \frac{p^2 + p_0^2}{2p_0^2} | \psi (\mathbf{p}) |^2 \, d\mathbf{p}$$

$$= \int | \psi (\mathbf{p}) |^2 \, d\mathbf{p} = 1 \qquad (6.14)$$

If for the sake of brevity we put

$$\lambda = \frac{Zme^2}{\hbar p_0} = \frac{Zme^2}{\hbar \, (- 2mE)^{1/2}} \qquad (6.15)$$

and go on to the new variables, the Schrödinger equation (6.8) becomes

$$\Psi (\alpha, \theta, \varphi) = \frac{\lambda}{2\pi^2} \int \frac{\Psi (\alpha', \theta', \varphi')}{4 \sin^2 (\omega/2)} \, d\Omega' \qquad (6.16)$$

Here $2 \sin (\omega/2)$ is the length of the chord and ω is the arc length of the great circle that connects the points α, θ, φ and α', θ', φ' on the four-dimensional sphere, so that

$$4 \sin^2 \frac{\omega}{2} = (\xi - \xi')^2 + (\eta - \eta')^2 + (\zeta - \zeta')^2 + (\chi - \chi')^2 \quad (6.17)$$

Equation (6.16) is the integral equation for the spherical harmonics of a four-dimensional sphere. The eigenvalues will be the integers $\lambda = n$ $(n = 1, 2, \ldots)$, and the eigenfunctions will be the homogeneous harmonic functions of degree $n - 1$ in its arguments ξ, η, ζ, χ, that is, functions of type

$$\Psi = u (\xi, \eta, \zeta, \chi) \qquad (6.18)$$

where $u(x_1, x_2, x_3, x_4)$ is the solution to the four-dimensional Laplace equation

$$\frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} + \frac{\partial^2 u}{\partial x_3^2} + \frac{\partial^2 u}{\partial x_4^2} = 0 \qquad (6.19)$$

As we see from (6.15) the integer n is the principal quantum number.

Thus the theory of the hydrogen atom is connected with the four-dimensional potential theory. This interrelation makes it possible to easily derive the properties of hydrogenic wave functions and, notably, establish the addition theorem for these functions. This theorem holds not only for real integral values of n (the discrete spectrum) but also for complex values (the continuous spectrum).

The most significant corollary of this interrelation is the determination of the *symmetry transformation group* allowed by the Schrödinger equation for the hydrogen atom. Equation (6.17) obviously retains its form under an orthogonal transformation of

the variables ξ, η, ζ, χ, that is, if the hypersphere is rotated arbitrarily in four-dimensional space. It follows from this that the initial Schrödinger equation not only possesses the usual spherical symmetry but also has a wider symmetry corresponding to four-dimensional rotations. This explains the long-known fact that the energy levels of hydrogen depend only on the principal quantum number n. Application of a wider transformation group to the Schrödinger equation ensures the results mentioned at the beginning of this section. We will not detail these results or their derivation.

Part V

DIRAC'S THEORY OF THE ELECTRON

Chapter I

THE DIRAC EQUATION

1. Quantum mechanics and the theory of relativity

The theories of both Schrödinger and Pauli are nonrelativistic. They ignore the fact that no mass point can move in space and no action can propagate in space at a velocity that exceeds the speed of light. A relativistic generalization of quantum mechanics requires introducing new physical concepts and even modifying the interpretation of the wave equation. The modification is necessary because we need to introduce, besides spin, a new degree of freedom for the electron and because we cannot interpret this degree of freedom within the limits of the one-body problem.

However, it is possible to formulate the problem of one body (the electron) in a given external electromagnetic field in accordance with the theory of relativity. Dirac formulated this problem when he suggested his equation for the electron.

In Section 13, Chapter III, Part I, we saw that the wave equation, that is, the equation that determines the law governing the time dependence of the electronic state (function ψ), must be

$$H\psi - i\hbar \frac{\partial \psi}{\partial t} = 0 \qquad (1.1)$$

where H is the Hamiltonian. The wave equation is closely related to the quantum equations of motion, from which we derived the wave equation (Section 13, Chapter III, Part I). The quantum equations in turn can be derived from the wave equation (Section 4, Chapter IV, Part I). Now, following Dirac, we must generalize the wave equation (1.1) to the theory of relativity. We must require that it be invariant under Lorentz trasnformations and that it give us the classical equations of motion of the theory of relativity.

2. Classical equations of motion

Let us recall the form of the classical equations of motion of the theory of relativity and the corresponding Lagrangian and Hamiltonian functions.

In the mechanics of the theory of relativity the momentum (P_x, P_y, P_z) is connected with the velocity $(\dot{x}, \dot{y}, \dot{z})$ by the relationships

$$P_x = \frac{m\dot{x}}{(1 - v^2/c^2)^{1/2}}, \quad P_y = \frac{m\dot{y}}{(1 - v^2/c^2)^{1/2}}, \quad P_z = \frac{m\dot{z}}{(1 - v^2/c^2)^{1/2}} \quad (2.1)$$

where

$$v^2 = \dot{x}^2 + \dot{y}^2 + \dot{z}^2 \tag{2.1*}$$

The equations of the motion of an electron (mass m and electric charge $-e$) in an electromagnetic field have the form

$$\frac{dP_x}{dt} = -\frac{e}{c}(\dot{y}\mathscr{H}_z - \dot{z}\mathscr{H}_y) - e\mathscr{E}_x$$

$$\frac{dP_y}{dt} = -\frac{e}{c}(\dot{z}\mathscr{H}_x - \dot{x}\mathscr{H}_z) - e\mathscr{E}_y$$

$$\frac{dP_z}{dt} = -\frac{e}{c}(\dot{x}\mathscr{H}_y - \dot{y}\mathscr{H}_x) - e\mathscr{E}_z \tag{2.2}$$

From these we can easily derive the equation

$$\frac{dT}{dt} = -e(\dot{x}\mathscr{E}_x + \dot{y}\mathscr{E}_y + \dot{z}\mathscr{E}_z) \tag{2.3}$$

where T is the electron's kinetic energy:

$$T = \frac{mc^2}{(1 - v^2/c^2)^{1/2}} \tag{2.4}$$

These equations can be obtained from the Lagrangian

$$\mathscr{L} = -mc^2(1 - v^2/c^2)^{1/2} - \frac{e}{c}(\dot{x}A_x + \dot{y}A_y + \dot{z}A_z) + e\Phi \tag{2.5}$$

where Φ is the scalar potential, and $\mathbf{A} = (A_x, A_y, A_z)$ the vector potential. The generalized momentum conjugate to coordinate x is

$$p_x = \frac{\partial \mathscr{L}}{\partial \dot{x}} = \frac{m\dot{x}}{(1 - v^2/c^2)^{1/2}} - \frac{e}{c}A_x = P_x - \frac{e}{c}A_x \tag{2.6}$$

and similarly for the other coordinates. Hence the generalized momenta p_x, p_y, p_z do not coincide with momentum components P_x, P_y, P_z but are linked with them, as in nonrelativistic case, by the relationships

$$P_x = p_x + \frac{e}{c}A_x, \quad P_y = p_y + \frac{e}{c}A_y, \quad P_z = p_z + \frac{e}{c}A_z \tag{2.7}$$

[see formula (5.16), Part III]. The energy of the electron is

$$E = \dot{x}p_x + \dot{y}p_y + \dot{z}p_z - \mathscr{L} = \frac{mc^2}{(1 - v^2/c^2)^{1/2}} - e\Phi \tag{2.8}$$

Expressing it in terms of the generalized momenta, we find the classical Hamiltonian function

$$H_{\text{classical}} = mc^2 \left[1 + \frac{1}{m^2c^2}\left(\mathbf{p} + \frac{e}{c}\,\mathbf{A}\right)^2\right]^{1/2} - e\Phi \qquad (2.9)$$

3. Derivation of the wave equation

We seek a quantum operator that corresponds to the Hamiltonian function (2.9). Let us start from the simplest case of a free electron, that is, when there is no electromagnetic field and the scalar and vector potentials are zero. Then

$$H_{\text{classical}} = mc^2 \left(1 + \frac{1}{m^2c^2}\left(p_x^2 + p_y^2 + p_z^2\right)\right)^{1/2} \qquad (3.1)$$

Because equations of the theory of relativity are symmetric with respect to coordinates and time and because the wave equation contains the linear operator of differentiation with respect to time, the equation must also have linear operators of differentiation with respect to coordinates. Consequently, the quantum Hamiltonian must be linear with respect to the operators

$$p_x = -\,i\hbar\,\frac{\partial}{\partial x}\,, \qquad p_y = -\,i\hbar\,\frac{\partial}{\partial y}\,, \qquad p_z = -\,i\hbar\,\frac{\partial}{\partial z}$$

that is, it must have the form

$$H = \beta_1 p_x + \beta_2 p_y + \beta_3 p_z + \beta_4 \qquad (3.2)$$

where β_k are the as yet unknown operators that do not depend on p_x, p_y, p_z. But these operators must not contain the coordinates x, y, z either because all points in space have equal status for a free electron. Consequently, they must act on some new variables on which the wave function in Schrödinger's theory did not depend. We will determine the meaning of these new variables later. We will see that they are a generalization of the operators in Pauli's theory.

To determine the properties of operators β_k we require that there be the same relationship between the square of the energy and the square of momentum of a free electron in quantum mechanics as in classical mechanics, namely

$$H^2 = m^2c^4 + c^2\left(p_x^2 + p_y^2 + p_z^2\right) \qquad (3.3)$$

Let us calculate the square of operator (3.2), bearing in mind that the β_k do not contain the coordinates and hence commute with p_x, p_y, p_z but cannot commute with each other. We get

$$H^2 = \beta_4^2 + \beta_1^2 p_x^2 + \beta_2^2 p_y^2 + \beta_3^2 p_z^2$$
$$+ \left(\beta_1\beta_4 + \beta_4\beta_1\right)p_x + \left(\beta_2\beta_4 + \beta_4\beta_2\right)p_y + \left(\beta_3\beta_4 + \beta_4\beta_3\right)p_z$$
$$+ \left(\beta_2\beta_3 + \beta_3\beta_2\right)p_y p_z + \left(\beta_3\beta_1 + \beta_1\beta_3\right)p_z p_x + \left(\beta_1\beta_2 + \beta_2\beta_1\right)p_x p_y \qquad (3.4)$$

This expression coincides with the previous one if the following conditions are observed:

$$\beta_4^2 = m^2c^4, \quad \beta_1^2 = \beta_2^2 = \beta_3^2 = c^2, \quad \beta_i\beta_k + \beta_k\beta_i = 0, \quad i \neq k \quad (3.5)$$

If we use the relationships

$$\beta_1 = c\alpha_1, \quad \beta_2 = c\alpha_2, \quad \beta_3 = c\alpha_3, \quad \beta_4 = mc^2\alpha_4 \quad (3.6)$$

to introduce new operators α_k proportional to β_k, then the Hamiltonian will be

$$H = c(\alpha_1 p_x + \alpha_2 p_y + \alpha_3 p_z) + mc^2\alpha_4 \quad (3.7)$$

and the α_k must satisfy the conditions

$$\alpha_k^2 = 1, \quad \alpha_i\alpha_k + \alpha_k\alpha_i = 0, \quad i \neq k \quad (3.8)$$

which in shorter form can be written as

$$\alpha_i\alpha_k + \alpha_k\alpha_i = 2\delta_{ik}, \quad i, k = 1, 2, 3, 4 \quad (3.9)$$

4. The Dirac matrices

We will see later that the α_k can be considered as transforming four functions ψ_1, ψ_2, ψ_3, ψ_4, in the same way that the Pauli matrices transform two. The object on which the α_k act will then be a collection of four functions, and the operators can be represented in matrix form with elements that are the coefficients of these transformations.

We will often denote the collection ψ_1, ψ_2, ψ_3, ψ_4 simply by ψ and the transformation

$$\psi_1' = \alpha_{11}\psi_1 + \alpha_{12}\psi_2 + \alpha_{13}\psi_3 + \alpha_{14}\psi_4$$
$$\psi_2' = \alpha_{21}\psi_1 + \alpha_{22}\psi_2 + \alpha_{23}\psi_3 + \alpha_{24}\psi_4$$
$$\psi_3' = \alpha_{31}\psi_1 + \alpha_{32}\psi_2 + \alpha_{33}\psi_3 + \alpha_{34}\psi_4$$
$$\psi_4' = \alpha_{41}\psi_1 + \alpha_{42}\psi_2 + \alpha_{43}\psi_3 + \alpha_{44}\psi_4 \quad (4.1)$$

by the brief notation

$$\psi' = \alpha\psi \quad (4.2)$$

where, consequently, α is the matrix

$$\alpha = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \alpha_{14} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} & \alpha_{24} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} & \alpha_{34} \\ \alpha_{41} & \alpha_{42} & \alpha_{43} & \alpha_{44} \end{pmatrix} \quad (4.3)$$

Let us express operators α_1, α_2, α_3, α_4, which satisfy (3.9), in terms of matrices similar to those in Pauli's theory (see Part III).

From α_1, α_2, α_3, α_4 we construct six matrices. First, three

$$\sigma_x = -i\alpha_2\alpha_3, \qquad \sigma_y = -i\alpha_3\alpha_1, \qquad \sigma_z = -i\alpha_1\alpha_2 \qquad (4.4)$$

and then, three more:

$$\rho_a = -i\alpha_1\alpha_2\alpha_3, \qquad \rho_b = \alpha_1\alpha_2\alpha_3\alpha_4, \qquad \rho_c = \alpha_4 \qquad (4.5)$$

We can easily verify that matrices σ_x, σ_y, σ_z will satisfy the same relationships as the Pauli matrices, namely

$$\sigma_y\sigma_z = -\sigma_z\sigma_y = i\sigma_x$$
$$\sigma_z\sigma_x = -\sigma_x\sigma_z = i\sigma_y$$
$$\sigma_x\sigma_y = -\sigma_y\sigma_z = i\sigma_z \qquad (4.6)$$

The square of each of these will be unity:

$$\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = 1 \qquad (4.7)$$

The matrices ρ_a, ρ_b, ρ_c will satisfy similar relationships:

$$\rho_b\rho_c = -\rho_c\rho_b = i\rho_a$$
$$\rho_c\rho_a = -\rho_a\rho_c = i\rho_b$$
$$\rho_a\rho_b = -\rho_b\rho_a = i\rho_c \qquad (4.8)$$

and

$$\rho_a^2 = \rho_b^2 = \rho_c^2 = 1 \qquad (4.9)$$

The products of matrices ρ into matrices σ are

$$\rho_a\sigma_x = \sigma_x\rho_a = \alpha_1$$
$$\rho_a\sigma_y = \sigma_y\rho_a = \alpha_2$$
$$\rho_a\sigma_z = \sigma_z\rho_a = \alpha_3 \qquad (4.10)$$

further

$$\rho_b\sigma_x = \sigma_x\rho_b = i\alpha_1\alpha_4$$
$$\rho_b\sigma_y = \sigma_y\rho_b = i\alpha_2\alpha_4$$
$$\rho_b\sigma_z = \sigma_z\rho_b = i\alpha_3\alpha_4 \qquad (4.11)$$

and, last

$$\rho_c\sigma_x = \sigma_x\rho_c = -i\alpha_2\alpha_3\alpha_4$$
$$\rho_c\sigma_y = \sigma_y\rho_c = -i\alpha_3\alpha_1\alpha_4$$
$$\rho_c\sigma_z = \sigma_z\rho_c = -i\alpha_1\alpha_2\alpha_4 \qquad (4.12)$$

Hence each of the matrices ρ commutes with each of the matrices σ, and we can say in a sense that ρ and σ refer to different degrees of freedom of the electron.

Using matrices α_i expressed in terms of ρ and σ, we can write the Hamiltonian (3.7) as

$$H = c\rho_a\,(\sigma_x p_x + \sigma_y p_y + \sigma_z p_z) + mc^2\rho_c \qquad (4.13)$$

We note that to the four matrices α_1, α_2, α_3, α_4, which satisfy (3.9), we could add a fifth, say, $\alpha_5 = \rho_b$. This last anticommutes with the Hamiltonian (4.13).

Now let us turn to the problem of building 4×4 matrices that have the general properties we have just stated. We first consider the three 2×2 matrices that we met in Pauli's theory.

If we denote them by σ_1^0, σ_2^0, σ_3^0, we have

$$\sigma_1^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2^0 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_3^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad (4.14)$$

Let us assume that the transformations

$$\sigma_1^0 \begin{pmatrix} \xi \\ \eta \end{pmatrix} = \begin{pmatrix} \eta \\ \xi \end{pmatrix}, \qquad \sigma_2^0 \begin{pmatrix} \xi \\ \eta \end{pmatrix} = \begin{pmatrix} -i\eta \\ i\xi \end{pmatrix}, \qquad \sigma_3^0 \begin{pmatrix} \xi \\ \eta \end{pmatrix} = \begin{pmatrix} \xi \\ -\eta \end{pmatrix} \qquad (4.15)$$

are applied not to one but to two pairs of numbers, $\begin{pmatrix} \xi \\ \eta \end{pmatrix}$ and $\begin{pmatrix} \xi^* \\ \eta^* \end{pmatrix}$, simultaneously. These two pairs can be considered to be a set of four numbers ψ_1, ψ_2, ψ_3, ψ_4. We can relate the numbers ξ, η, ξ^*, η^* to ψ_1, ψ_2, ψ_3, ψ_4 in different ways.

We can put, for instance

$$\psi_1 = \xi, \qquad \psi_2 = \eta, \qquad \psi_3 = \xi^*, \qquad \psi_4 = \eta^* \qquad (4.16)$$

or we can put

$$\psi_1 = \xi, \qquad \psi_2 = \xi^*, \qquad \psi_3 = \eta, \qquad \psi_4 = \eta^* \qquad (4.17)$$

In the first case the matrices corresponding to our transformations will be

$$\sigma_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$
$$(4.18)$$

and in the second case

$$\rho_1 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \quad \rho_2 = \begin{pmatrix} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & -i \\ i & 0 & 0 & 0 \\ 0 & i & 0 & 0 \end{pmatrix}, \quad \rho_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$
$$(4.19)$$

Obviously, the transformations σ_1, σ_2, σ_3 and the transformations ρ_1, ρ_2, ρ_3, taken separately, satisfy the same relationships as the transformations (4.15), which are applied to two functions. Namely

$$\sigma_2\sigma_3 = -\sigma_3\sigma_2 = i\sigma_1$$
$$\sigma_3\sigma_1 = -\sigma_1\sigma_3 = i\sigma_2$$
$$\sigma_1\sigma_2 = -\sigma_2\sigma_1 = i\sigma_3$$
$$\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = 1 \tag{4.20}$$

and

$$\rho_2\rho_3 = -\rho_3\rho_2 = i\rho_1$$
$$\rho_3\rho_1 = -\rho_1\rho_3 = i\rho_2$$
$$\rho_1\rho_2 = -\rho_2\rho_1 = i\rho_3$$
$$\rho_1^2 = \rho_2^2 = \rho_3^2 = 1 \tag{4.21}$$

On the other hand, we can verify that each transformation σ commutes with each transformation ρ, so that

$$\sigma_i\rho_k = \rho_k\sigma_i, \qquad i,\ k = 1,\ 2,\ 3 \tag{4.22}$$

Each of the matrices ρ and σ and each of their products (4.22) has two 2-fold degenerate eigenvalues, $+1$ and -1 .

The three matrices σ_i, the three ρ_i, and the nine $\sigma_i\rho_k$ together with the unit matrix form a system of 16 matrices, which can be called complete in the sense that any 4×4 matrix, that is, a matrix with 16 elements, can be expressed as a linear combination of these 16 matrices. The coefficients in this combination are numbers.

For example, we can use ρ_i and σ_i to express the matrices involved in the Dirac equation and the related matrices, ρ_a, ρ_b, ρ_c and σ_x, σ_y, σ_z. We can do this in various ways, so that matrices with a given physical meaning can have different mathematical forms. In the literature we most often find the representation introduced by Dirac, who put

$$\sigma_x = \sigma_1, \quad \sigma_y = \sigma_2, \quad \sigma_z = \sigma_3, \quad \rho_a = \rho_1, \quad \rho_b = \rho_2, \quad \rho_c = \rho_3 \tag{4.23}$$

According to (4.10) and (4.15), the corresponding matrices α_k will be

$$\alpha_1' = \rho_1\sigma_1, \qquad \alpha_2' = \rho_1\sigma_2, \qquad \alpha_3' = \rho_1\sigma_3, \qquad \alpha_4' = \rho_3 \tag{4.24}$$

(We have assigned a prime for each matrix so as to distinguish these from the ones that we will use later.)

In some respects it is more convenient to use the following matrices:

$$\sigma_x = \rho_3\sigma_1, \qquad \sigma_y = \sigma_2, \qquad \sigma_z = \rho_3\sigma_3$$
$$\rho_a = \rho_3, \qquad \rho_b = \rho_1\sigma_2, \qquad \rho_c = \rho_2\sigma_2 \tag{4.25}$$

or in explicit form

$$\sigma_x = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

(4.26)

$$\rho_a = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \quad \rho_b = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix}, \quad \rho_c = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}$$

(4.27)

From this we get the following expressions for α_k:

$$\alpha_1 = \sigma_1, \qquad \alpha_2 = \rho_3\sigma_2, \qquad \alpha_3 = \sigma_3, \qquad \alpha_4 = \rho_2\sigma_2 \qquad (4.28)$$

or in explicit form

$$\alpha_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad \alpha_2 = \begin{pmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & i \\ 0 & 0 & -i & 0 \end{pmatrix},$$

$$\alpha_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \quad \alpha_4 = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} \qquad (4.29)$$

5. The Dirac equation for a free electron

We can now write the Dirac equation for a free electron in explicit form. If H is the operator (3.7), the wave equation

$$H\psi = [c\,(\alpha_1 p_x + \alpha_2 p_y + \alpha_3 p_z) + mc^2\alpha_4]\,\psi = i\hbar \frac{\partial\psi}{\partial t} \qquad (5.1)$$

can be written as a system of four differential equations:

$$-i\hbar c \left(\frac{\partial\psi_2}{\partial x} - i\,\frac{\partial\psi_2}{\partial y} + \frac{\partial\psi_1}{\partial z}\right) - mc^2\psi_4 = i\hbar\,\frac{\partial\psi_1}{\partial t}$$

$$-i\hbar c \left(\frac{\partial\psi_1}{\partial x} + i\,\frac{\partial\psi_1}{\partial y} - \frac{\partial\psi_2}{\partial z}\right) + mc^2\psi_3 = i\hbar\,\frac{\partial\psi_2}{\partial t}$$

$$-i\hbar c \left(\frac{\partial\psi_4}{\partial x} + i\,\frac{\partial\psi_4}{\partial y} + \frac{\partial\psi_3}{\partial z}\right) + mc^2\psi_2 = i\hbar\,\frac{\partial\psi_3}{\partial t}$$

$$-i\hbar c \left(\frac{\partial\psi_3}{\partial x} - i\,\frac{\partial\psi_3}{\partial y} - \frac{\partial\psi_4}{\partial z}\right) - mc^2\psi_1 = i\hbar\,\frac{\partial\psi_4}{\partial t} \qquad (5.2)$$

We note that it is easier to investigate the Dirac equation when it is written in the short form (5.1), so that there will be almost no need to use (5.2).

We have considered two ways of choosing matrices. One, proposed by Dirac, corresponds to formula (4.23). The other, proposed in this book, corresponds to (4.25).

For some purposes it is convenient to introduce such a representation of matrices α_k as will ensure that the corresponding system of equations for the four component wave function of the free electron has real coefficients. It is sufficient to interchange ρ_2 and ρ_3 in (4.25) and change the sign before matrix ρ_1. Instead of (4.25) we will then have

$$\sigma_x^0 = \rho_2\sigma_1, \qquad \sigma_y^0 = \sigma_2, \qquad \sigma_z^0 = \rho_2\sigma_1$$
$$\rho_a^0 = \rho_2, \qquad \rho_b = -\rho_1\sigma_2, \qquad \rho_c^0 = \rho_3\sigma_2 \qquad (5.3)$$

To distinguish the new matrices from the old we assigned them the index 0. All matrices (5.3) have pure imaginary elements. [1]

The new matrices and the old are connected via a canonical transformation with matrix

$$T = \frac{1}{\sqrt{2}}(\rho_2 + \rho_3) \qquad (5.4)$$

which is hermitian and unitary, so that

$$T^{-1} = T^+ = T, \qquad T^2 = 1 \qquad (5.5)$$

Indeed, we have

$$T\rho_2 T = \rho_3, \qquad T\rho_3 T = \rho_2, \qquad T\rho_1 T = -\rho_1 \qquad (5.6)$$

The new matrices α_k (which we denote by α_k^0) will be connected with the old matrices through the canonical transformation

$$\alpha_k^0 = T^+\alpha_k T \qquad (5.7)$$

These will be

$$\alpha_1^0 = \sigma_1, \qquad \alpha_2^0 = \rho_2\sigma_2, \qquad \alpha_3^0 = \sigma_3, \qquad \alpha_4^0 = \rho_3\sigma_2 \qquad (5.8)$$

A comparison with (4.28) shows that they differ from the old in that α_2 is interchanged with α_4. The elements of the first three matrices α_k^0 will be real, and the elements of α_4^0 pure imaginary.

---

[1] Since we are dealing with 4×4 matrices, there is no danger of confusing them with the Pauli matrices (4.14).

We can now write the system of four differential equations for the wave function of a free electron, (5.2), as

$$\frac{\partial \psi_2^0}{\partial x} - \frac{\partial \psi_4^0}{\partial y} + \frac{\partial \psi_1^0}{\partial z} + \frac{1}{c} \frac{\partial \psi_1^0}{\partial t} + \frac{mc}{\hbar} \psi_2^0 = 0$$

$$\frac{\partial \psi_1^0}{\partial x} + \frac{\partial \psi_3^0}{\partial y} - \frac{\partial \psi_2^0}{\partial z} + \frac{1}{c} \frac{\partial \psi_2^0}{\partial t} - \frac{mc}{\hbar} \psi_1^0 = 0$$

$$\frac{\partial \psi_4^0}{\partial x} + \frac{\partial \psi_2^0}{\partial y} + \frac{\partial \psi_3^0}{\partial z} + \frac{1}{c} \frac{\partial \psi_3^0}{\partial t} - \frac{mc}{\hbar} \psi_4^0 = 0$$

$$\frac{\partial \psi_3^0}{\partial x} - \frac{\partial \psi_1^0}{\partial y} - \frac{\partial \psi_4^0}{\partial z} + \frac{1}{c} \frac{\partial \psi_4^0}{\partial t} + \frac{mc}{\hbar} \psi_3^0 = 0 \qquad (5.9)$$

Last, let us write in explicit form the relationship between the wave functions ψ_k', which correspond to the choice of matrices α_k' according to Dirac [see 4.24], and our wave functions. We have

$$\psi_1' = \frac{\psi_1 - \psi_4}{\sqrt{2}}, \qquad \psi_2' = \frac{\psi_2 + \psi_3}{\sqrt{2}}$$

$$\psi_3' = \frac{\psi_1 + \psi_4}{\sqrt{2}}, \qquad \psi_4' = \frac{\psi_2 - \psi_3}{\sqrt{2}} \qquad (5.10)$$

We introduce the unitary matrix S that corresponds to the transformation

$$\psi' = S\psi \qquad (5.11)$$

This matrix is

$$S = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & -1 & 0 \end{pmatrix} = \frac{1 - i\rho_2}{\sqrt{2}} \frac{1 + i\sigma_2}{\sqrt{2}} \frac{1 - i\rho_3\sigma_2}{\sqrt{2}} \qquad (5.12)$$

As for the relationship between ψ^0 and ψ, which corresponds to the transformation

$$\psi^0 = T\psi \qquad (5.13)$$

this is given by the formulas

$$\psi_1^0 = \frac{1}{\sqrt{2}} (\psi_1 - i\psi_3), \qquad \psi_2^0 = \frac{1}{\sqrt{2}} (\psi_2 - i\psi_4)$$

$$\psi_3^0 = \frac{1}{\sqrt{2}} (- \psi_3 + i\psi_1), \qquad \psi_4^0 = \frac{1}{\sqrt{2}} (- \psi_4 + i\psi_2) \qquad (5.14)$$

Since $T^2 = 1$, the same formulas hold for ψ_i and ψ_i^0. We have

$$\psi_1 = \frac{1}{\sqrt{2}} (\psi_1^0 - i\psi_3^0), \qquad \psi_2 = \frac{1}{\sqrt{2}} (\psi_2^0 - i\psi_4^0)$$

$$\psi_3 = \frac{1}{\sqrt{2}} (- \psi_3^0 + i\psi_1^0), \qquad \psi_4 = \frac{1}{\sqrt{2}} (- \psi_4^0 + i\psi_2^0) \qquad (5.15)$$

6. Lorentz transformations

We will now prove the invariance of the wave equation under Lorentz transformations and investigate the geometric properties of ψ_1, ψ_2, ψ_3, ψ_4.

We put

$$x = x_1, \qquad y = x_2, \qquad z = x_3, \qquad ct = x_0 \tag{6.1}$$

and introduce four numbers:

$$e_0 = 1, \qquad e_1 = e_2 = e_3 = -1 \tag{6.2}$$

so that we can write the square of the four-dimensional *separation* in the form

$$\pm ds^2 = c^2 \, dt^2 - dx^2 - dy^2 - dz^2 = \sum_{k=0}^{3} e_k \, dx_k^2 \tag{6.3}$$

We write the Lorentz transformations as

$$x_i' = \sum_{k=0}^{3} e_k a_{ik} x_k \tag{6.4}$$

where a_{ik} are real numbers that satisfy the condition

$$\sum_{i=0}^{3} e_i a_{ik} a_{il} = e_k \delta_{kl} \tag{6.5}$$

These conditions follow from the fact that transformations (6.4) must leave ds^2 invariant. Owing to these conditions the solution of (6.4) for x_i yields

$$x_i = \sum_{k=0}^{3} e_k a_{ki} x_k' \tag{6.6}$$

which in turn leads to the equations

$$\sum_{i=0}^{3} e_i a_{ki} a_{li} = e_k \delta_{kl} \tag{6.7}$$

If we multiply Eq. (5.1) by $i/(\hbar c)$, we can write it in the form

$$\sum_{k=1}^{3} \alpha_k \frac{\partial \psi}{\partial x_k} + \frac{imc}{\hbar} \alpha_4 \psi + \frac{\partial \psi}{\partial x_0} = 0 \tag{6.8}$$

or

$$\sum_{k=0}^{3} \alpha_k \frac{\partial \psi}{\partial x_k} + \frac{imc}{\hbar} \alpha_4 \psi = 0 \tag{6.9}$$

provided that α_0 is the unit matrix.

19*

Now let us change variables according to (6.6). We have

$$\frac{\partial \psi}{\partial x_k} = \sum_{l=0}^{3} e_k a_{lk} \frac{\partial \psi}{\partial x_l'}$$

Hence

$$\sum_{l=0}^{3} \sum_{k=0}^{3} e_k a_{lk} \alpha_k \frac{\partial \psi}{\partial x_l'} + \frac{imc}{\hbar} \alpha_4 \psi = 0 \tag{6.10}$$

If we can find a matrix S (not a unitary one, generally speaking) such that

$$\alpha_l' = S^+ \alpha_l S = \sum_{k=0}^{3} e_k a_{lk} \alpha_k, \qquad l = 0, 1, 2, 3 \tag{6.11}$$

and

$$S^+ \alpha_4 S = \alpha_4 \tag{6.12}$$

then Eq. (6.10) can be written thus:

$$\sum_{l=0}^{3} S^+ \alpha_l S \frac{\partial \psi}{\partial x_l'} + \frac{imc}{\hbar} S^+ \alpha_4 S \psi = 0 \tag{6.13}$$

If we then put

$$\psi' = S\psi \tag{6.14}$$

and premultiply (6.13) by $(S^+)^{-1}$, that is, initiate a transformation of the four equations (6.13) inverse to S^+, we get

$$\sum_{k=0}^{3} \alpha_k \frac{\partial \psi'}{\partial x_k'} + \frac{imc}{\hbar} \alpha_4 \psi' = 0 \tag{6.15}$$

that is, an equation of the same type as the initial one, (6.9), with the same matrices α_k but with new independent variables x_0', x_1', x_2', x_3' and new functions ψ_1', ψ_2', ψ_3', ψ_4'. Hence we will be able to prove that

If the Lorentz transformations are accompanied by transformation (6.14), which acts on functions ψ, the wave equation will keep its form.

In other words, we will prove the invariance of the wave equation under the Lorentz transformations.

7. Form of matrix S for spatial rotations of axes and for Lorentz transformations

We will show that a matrix S with the needed properties does indeed exist and that with our choice of matrices α_k it has the form

$$S = \begin{pmatrix} \alpha & \beta & 0 & 0 \\ \gamma & \delta & 0 & 0 \\ 0 & 0 & \bar{\alpha} & \bar{\beta} \\ 0 & 0 & \bar{\gamma} & \bar{\delta} \end{pmatrix} \tag{7.1}$$

where α, β, γ, δ are four complex parameters connected by the relationship

$$\alpha\delta - \beta\gamma = 1 \tag{7.2}$$

These parameters are called the *generalized Cayley-Klein parameters*. The hermitian conjugate of S will be

$$S^+ = \begin{pmatrix} \bar{\alpha} & \bar{\gamma} & 0 & 0 \\ \bar{\beta} & \bar{\delta} & 0 & 0 \\ 0 & 0 & \alpha & \gamma \\ 0 & 0 & \beta & \delta \end{pmatrix} \tag{7.3}$$

First of all it is easy to prove by direct calculation that Eq. (6.12) is an identity owing to the relationship (7.2) if we perform three transformations successively: first S, then α_4, and last S^+. Note that we also have

$$S^+ \alpha_5 S = \alpha_5 \tag{7.4}$$

To make certain that for any Lorentz transformation we can choose parameters α, β, γ, δ so that Eq. (6.11) will hold, we will use the fact that the Lorentz transformations and the transformations S each form a group, that is, that several successive transformations can be replaced by one of the same type. We can obtain the most general Lorentz transformation equations by applying successively transformations that are simple in form. For instance, we can rotate the coordinate system about the x, y, z axes and then apply the transformation

$$z' = \frac{z - vt}{(1 - v^2/c^2)^{1/2}}, \qquad t' = \frac{t - vz/c^2}{(1 - v^2/c^2)^{1/2}} \tag{7.5}$$

If we then find the matrix S for each of these transformations, the matrix S of the general transformation can be obtained by multiplying the matrices of all the transformations.

Let us first study the rotation about the z axis, since in this case S has the simplest form. We write the transformation equations as

$$x_1' = x_1 \cos \varphi - x_2 \sin \varphi$$
$$x_2' = x_1 \sin \varphi + x_2 \cos \varphi$$
$$x_3' = x_3$$
$$x_0' = x_0 \tag{7.6}$$

We will show that the parameters corresponding to this rotation are

$$\alpha = e^{-i\varphi/2}, \qquad \beta = 0, \qquad \gamma = 0, \qquad \delta = e^{i\varphi/2} \tag{7.7}$$

so that the matrix of this transformation is

$$S = \begin{pmatrix} e^{-i\varphi/2} & 0 & 0 & 0 \\ 0 & e^{i\varphi/2} & 0 & 0 \\ 0 & 0 & e^{i\varphi/2} & 0 \\ 0 & 0 & 0 & e^{-i\varphi/2} \end{pmatrix} \tag{7.8}$$

We can write it also as

$$S = \cos \frac{\varphi}{2} - i \sin \frac{\varphi}{2} \rho_3 \sigma_3 \tag{7.9}$$

or, according to (4.25), as

$$S = \cos \frac{\varphi}{2} - i \sin \frac{\varphi}{2} \sigma_z \tag{7.10}$$

We have

$$S^+ \alpha_1 S = S^+ \rho_a \sigma_x S = \rho_a \left(\cos \frac{\varphi}{2} + i \sin \frac{\varphi}{2} \sigma_z \right) \sigma_x \left(\cos \frac{\varphi}{2} - i \sin \frac{\varphi}{2} \sigma_z \right)$$
$$= \rho_a (\cos \varphi + i \sin \varphi \sigma_z) \sigma_x = \rho_a (\sigma_x \cos \varphi - \sigma_y \sin \varphi)$$

so that

$$\alpha_1' = S^+ \alpha_1 S = \alpha_1 \cos \varphi - \alpha_2 \sin \varphi \tag{7.11}$$

and in a similar manner

$$\alpha_2' = S^+ \alpha_2 S = \alpha_1 \sin \varphi + \alpha_2 \cos \varphi$$
$$\alpha_3' = S^+ \alpha_3 S = \alpha_3$$
$$\alpha_0' = S^+ \alpha_0 S = \alpha_0 \tag{7.11*}$$

Thus the new matrices α_k' are expressed in terms of the old α_k in the same way as the new (transformed) coordinates x_k' are expressed in terms of the old x_k, which means that relationships (6.11) hold true.

Now we study the rotation about the axis $x = x_1$:

$$x'_1 = x_1$$
$$x'_2 = x_2 \cos \varphi - x_3 \sin \varphi$$
$$x'_3 = x_2 \sin \varphi + x_3 \cos \varphi$$
$$x'_0 = x_0 \tag{7.12}$$

Here

$$\alpha = \cos \frac{\varphi}{2}, \quad \beta = -i \sin \frac{\varphi}{2}, \quad \gamma = -i \sin \frac{\varphi}{2}, \quad \delta = \cos \frac{\varphi}{2} \tag{7.13}$$

so that the transformation matrix

$$S = \cos \frac{\varphi}{2} - i \sin \frac{\varphi}{2} \rho_3 \sigma_1 \tag{7.14}$$

or

$$S = \cos \frac{\varphi}{2} - i \sin \frac{\varphi}{2} \sigma_x \tag{7.14*}$$

To prove that (6.11) is valid in this case we can use the same method as in the previous case.

Finally, for the rotation about the axis $y = x_2$ we have

$$x'_1 = x_1 \cos \varphi + x_3 \sin \varphi$$
$$x'_2 = x_2$$
$$x'_3 = -x_1 \sin \varphi + x_3 \cos \varphi$$
$$x'_0 = x_0 \tag{7.15}$$

The parameters in this case are

$$\alpha = \cos \frac{\varphi}{2}, \quad \beta = -\sin \frac{\varphi}{2}, \quad \gamma = \sin \frac{\varphi}{2}, \quad \delta = \cos \frac{\varphi}{2} \tag{7.16}$$

If we use these parameters to build matrix S for this rotation, we get

$$S = \cos \frac{\varphi}{2} - i \sin \frac{\varphi}{2} \sigma_2 \tag{7.17}$$

or

$$S = \cos \frac{\varphi}{2} - i \sin \frac{\varphi}{2} \sigma_y \tag{7.17*}$$

In all three cases a rotation about axis x_k by an angle φ in the positive direction has corresponding to it a unitary matrix

$$S = \cos \frac{\varphi}{2} - i \sin \frac{\varphi}{2} \sigma_{x_k} \tag{7.18}$$

where the result does not depend on the choice of matrices α_k. To generalize the result we can say that a spatial rotation by an

angle ω about an axis with direction cosines l, m, n has corresponding to it a matrix

$$S = \cos\frac{\omega}{2} - i\sin\frac{\omega}{2}(l\sigma_x + m\sigma_y + n\sigma_z) \qquad (7.19)$$

We now turn to the Lorentz transformation equations proper, (7.5), which we write as

$$x_1' = x_1, \qquad\qquad x_2' = x_2$$
$$x_3' = \frac{x_3 - vx_0/c}{(1 - v^2/c^2)^{1/2}}, \qquad x_0' = \frac{x_0 - vx_3/c}{(1 - v^2/c^2)^{1/2}} \qquad (7.20)$$

We put

$$\frac{v}{c} = \tanh u \qquad (7.21)$$

so that

$$(1 - v^2/c^2)^{-1/2} = \cosh u, \qquad (v/c)(1 - v^2/c^2)^{-1/2} = \sinh u \quad (7.21^*)$$

The parameters in this case are

$$\alpha = e^{-u/2}, \qquad \beta = 0, \qquad \gamma = 0, \qquad \delta = e^{u/2} \qquad (7.22)$$

Here matrix S will no longer be unitary but with our choice of matrices α_k will remain of type (7.1). We write it as

$$S = \cosh\frac{u}{2} - \sinh\frac{u}{2}\sigma_3 \qquad (7.23)$$

or

$$S = \cosh\frac{u}{2} - \sinh\frac{u}{2}\alpha_3 \qquad (7.24)$$

It is easy to verify that now too the relationships (6.11) hold.

For the Lorentz transformation corresponding to the motion along the x_k axis with the velocity $v = c\tanh u$ the matrix S is

$$S = \cosh\frac{u}{2} - \sinh\frac{u}{2}\alpha_k \qquad (7.25)$$

and for the transformation corresponding to the motion along the line with direction cosines l, m, n

$$S = \cosh\frac{u}{2} - \sinh\frac{u}{2}(l\alpha_1 + m\alpha_2 + n\alpha_3) \qquad (7.26)$$

Hence in all cases we can find a matrix S of type (7.1) that satisfies relationships (6.11). We have thus proved the invariance of the wave equation under a Lorentz transformation.

We note that the Cayley-Klein parameters and hence matrix S are determined for a given rotation to within their sign. In our formulas the sign of matrix S has been chosen so that an infinitesimal rotation, or a Lorentz transformation with an infinitesimal velocity, has corresponding to it a matrix that differs from $S = +1$ by a second-order infinitesimal.

8. Current density

We see that to each Lorentz transformation there corresponds a definite (to within its sign) transformation of the functions ψ_1, ψ_2, ψ_3, ψ_4. These functions therefore represent a kind of geometric quantity, such as a vector or tensor. We can call this quantity a *tensor of rank* $1/2$ or *half-vector*. The justification for this name is that some quadratic forms of ψ are transformed as a four-dimensional vector. Indeed, let us put

$$A_k = \bar{\psi} \alpha_k \psi, \qquad k = 0, 1, 2, 3, 4, 5 \tag{8.1}$$

or in explicit form (with our choice of α_k)

$$A_0 = \bar{\psi}_1 \psi_1 + \bar{\psi}_2 \psi_2 + \bar{\psi}_3 \psi_3 + \bar{\psi}_4 \psi_4$$

$$A_1 = \bar{\psi}_1 \psi_2 + \bar{\psi}_2 \psi_1 + \bar{\psi}_3 \psi_4 + \bar{\psi}_4 \psi_3$$

$$A_2 = - i\bar{\psi}_1 \psi_2 + i\bar{\psi}_2 \psi_1 + i\bar{\psi}_3 \psi_4 - i\bar{\psi}_4 \psi_3$$

$$A_3 = \bar{\psi}_1 \psi_1 - \bar{\psi}_2 \psi_2 + \bar{\psi}_3 \psi_3 - \bar{\psi}_4 \psi_4$$

$$A_4 = - \bar{\psi}_1 \psi_4 + \bar{\psi}_2 \psi_3 + \bar{\psi}_3 \psi_2 - \bar{\psi}_4 \psi_1$$

$$A_5 = - i\bar{\psi}_1 \psi_4 + i\bar{\psi}_2 \psi_3 - i\bar{\psi}_3 \psi_2 + i\bar{\psi}_4 \psi_1 \tag{8.2}$$

Formulas (6.11) then give

$$A'_l = \bar{\psi}' \alpha_l \psi' = \bar{\psi} S^+ \alpha_l S \psi = \bar{\psi} \alpha'_l \psi = \sum_{k=0}^{3} e_k a_{kl} A_k, \qquad l = 0, 1, 2, 3 \tag{8.3}$$

These equalities show that A_0, A_1, A_2, A_3 transform like the components of a four-dimensional vector. The formulas (6.12) and (7.4), on the other hand, give

$$A'_4 = A_4, \qquad A'_5 = A_5 \tag{8.4}$$

which means that the quantities A_4 and A_5 are four-dimensional invariants. The quantities A_k are linked thus:

$$A_1^2 + A_2^2 + A_3^2 + A_4^2 + A_5^2 = A_0^2 \tag{8.5}$$

Let us show that if the A_k are built by means of the functions ψ, which satisfy the wave equation, we have the following equation:

$$\frac{\partial A_1}{\partial x} + \frac{\partial A_2}{\partial y} + \frac{\partial A_3}{\partial z} + \frac{1}{c} \frac{\partial A_0}{\partial t} = 0 \tag{8.6}$$

For this we write Eq. (6.9) and its complex conjugate:

$$\sum_{k=0}^{3} \alpha_k \frac{\partial \psi}{\partial x_k} + \frac{imc}{\hbar} \alpha_4 \psi = 0 \tag{8.7}$$

$$\sum_{k=0}^{3} \frac{\partial \bar{\psi}}{\partial x_k} \alpha_k - i \frac{mc}{\hbar} \bar{\psi} \alpha_4 = 0 \tag{8.7*}$$

We premultiply the first of these equations by $\bar{\psi}$ and the second we multiply into ψ, and we add the results. The second terms in the equations cancel out, and we get

$$\sum_{k=0}^{3} \frac{\partial}{\partial x_k} (\bar{\psi} \alpha_k \psi) = 0 \qquad (8.8)$$

that is, we come to Eq. (8.6)

If we now integrate (8.6) over a volume V encompassed by a surface σ, we get

$$\frac{\partial}{\partial t} \int_V A_0 \, d\tau = - c \int [A_1 \cos(n, x) + A_2 \cos(n, y) + A_3 \cos(n, z)] \, d\tau \qquad (8.9)$$

The physical meaning of A_0 is the probability density. In Schrödinger's theory this is $\bar{\psi}\psi$. A comparison of (8.6) and (8.9) with (2.7) and (2.9) of Chapter III, Part II, shows that vector cA_k plays the role of vector $\mathbf{S}$ in Schrödinger's theory and is the counterpart of the electron flux. Hence the classical charge density ρ and current density ρv have corresponding to them the quantum counterparts

$$\rho \to - e \bar{\psi} \psi = - e A_0 \qquad (8.10)$$

$$\rho v_k \to - e S_k = - e c \bar{\psi} \alpha_k \psi = - e c A_k, \qquad k = 1, \ 2, \ 3 \qquad (8.11)$$

9. The Dirac equation in the case of a field.
Equations of motion

The wave equation for the free electron introduced in previous sections had the form

$$H\psi - i\hbar \frac{\partial \psi}{\partial t} = 0 \qquad (9.1)$$

with the Hamiltonian

$$H = c \left(\alpha_1 p_x + \alpha_2 p_y + \alpha_3 p_z \right) + mc^2 \alpha_4 \qquad (9.2)$$

We must now generalize this equation for an electron in an electromagnetic field. The classical Hamiltonian function, (2.9), can be found from the function in the case of no field, provided that we add to the latter the potential energy $-e\varphi$ and change the generalized momenta p_x, p_y, p_z to momenta P_x, P_y, P_z according to (2.7) We did this in Pauli's theory. Let us try the same substitution in our relativistic Hamiltonian (9.2) and put

$$H = c \left[\alpha_1 P_x + \alpha_2 P_y + \alpha_3 P_z \right] + mc^2 \alpha_4 - e\varphi \qquad (9.8)$$

where

$$P_x = p_x + \frac{e}{c} A_x, \quad P_y = p_y + \frac{e}{c} A_y, \quad P_z = p_z + \frac{e}{c} A_z \qquad (9.4)$$

and where p_x, p_y, p_z are understood to be the operators

$$-i\hbar \frac{\partial}{\partial x}, \qquad -i\hbar \frac{\partial}{\partial y}, \qquad -i\hbar \frac{\partial}{\partial z}$$

The main justification for this transition from the equation without a field to the equation for an electron in an electromagnetic field is this. The directly observable physical quantities are the electric and magnetic fields $\mathscr{E}$ and $\mathscr{H}$. The potentials, on the other hand, are auxiliary mathematical functions determined only to within the transformations

$$\mathbf{A}' = \mathbf{A} + \operatorname{grad} f, \qquad \varphi' = \varphi - \frac{1}{c} \frac{\partial f}{\partial t} \tag{9.5}$$

which leave the field unchanged. For this reason we must require that all physical effects that follow from the wave equation remain unchanged when we substitute $\mathbf{A}'$ and φ' for $\mathbf{A}$ and φ. This requirement will be met if such a substitution has corresponding to it a unitary transformation of operators and wave functions. Let us show that this proposition is true if the Hamiltonian is of type (9.3).

Let us denote by H' an operator of type (9.3) in which $\mathbf{A}'$ and φ' are substituted for $\mathbf{A}$ and φ by means of (9.5). If we use relations of type

$$\left[-i\hbar \frac{\partial}{\partial x} + \frac{e}{c}\left(\mathscr{A}_x + \frac{\partial f}{\partial x}\right)\right]\psi' = e^{-\frac{ie}{\hbar c}f}\left[-i\hbar \frac{\partial}{\partial x} + \frac{e}{c}\mathscr{A}_n\right]e^{\frac{ie}{\hbar c}f}\psi' \tag{9.6}$$

we can easily show that, if ψ satisfies (9.1), the function

$$\psi' = e^{-ief/\hbar c}\psi \tag{9.7}$$

will be the solution to

$$H'\psi' - i\hbar \frac{\partial \psi'}{\partial t} = 0 \tag{9.8}$$

Thus adding a gradient to the vector potential is equivalent to introducing a phase factor into the wave function, which is a particular case of a unitary transformation.

The wave equation (9.1) with the Hamiltonian (9.3) will obviously be invariant under a Lorentz transformation because the vector potential is transformed in the same way as the gradient, and the scalar potential as $-(1/c)(\partial/\partial t)$. Besides, and this can easily be shown, Eq. (6.8) is still correct.

How can we prove the validity of the new wave equation? A formal proof is to consider the equations of motion. We recall the formula for the total time derivative of an operator L [see (13.22), Chapter III, Part I], that is,

$$\frac{dL}{dt} = \frac{\partial L}{\partial t} + \frac{i}{\hbar}(HL - LH) \tag{9.9}$$

Let us see whether substituting x, y, z, P_x, P_y, P_z successively for L brings us to the classical equations of motion, which was the case in Schrödinger's theory. We put

$$L = x, \; y, \; z$$

and obtain

$$\dot{x} = \frac{dx}{dt} = c\alpha_1, \qquad \dot{y} = \frac{dy}{dt} = c\alpha_2, \qquad \dot{z} = \frac{dz}{dt} = c\alpha_3 \quad (9.10)$$

These are the operators for the velocity components of the electron. They do not commute with each other, and the square of each is c^2, that is, the square of the speed of light. The eigenvalues of the three operators are $\pm c$. It turns out then that measurement of a component of the velocity of an electron is certain to lead to the result $\pm c$. The question of whether this paradoxical result has a physical meaning remains open. The author is inclined to consider this result as a defect of Dirac's theory.

Let us now put

$$L = P_x = p_x + \frac{e}{c} A_x$$

and then calculate dP_x/dt. For this we first find

$$\frac{i}{\hbar}(P_y P_z - P_z P_y) = \frac{e}{c}\left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}\right) = \frac{e}{c}\mathscr{H}_x$$

$$\frac{i}{\hbar}(P_z P_x - P_x P_z) = \frac{e}{c}\left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}\right) = \frac{e}{c}\mathscr{H}_y$$

$$\frac{i}{\hbar}(P_x P_y - P_y P_x) = \frac{e}{c}\left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right) = \frac{e}{c}\mathscr{H}_z \qquad (9.11)$$

On the basis of this we have

$$\frac{dP_x}{dt} = \frac{e}{c}\frac{\partial A_x}{\partial t} + \frac{ic}{\hbar}\left[\alpha_2(P_y P_x - P_x P_y) + \alpha_3(P_z P_x - P_x P_z)\right]$$

$$\qquad - \frac{ie}{\hbar}(\varphi P_x - P_x \varphi)$$

$$= e\left(\frac{1}{c}\frac{\partial A_x}{\partial t} + \frac{\partial \varphi}{\partial x}\right) - e\alpha_2\mathscr{H}_z + e\alpha_3\mathscr{H}_y$$

or

$$\frac{dP_x}{dt} = - e(\alpha_2\mathscr{H}_z - \alpha_3\mathscr{H}_y) - e\mathscr{E}_x \qquad (9.12)$$

The last step is to introduce by means of (9.10) the operators $\dot{x}$, $\dot{y}$, $\dot{z}$ and write for dP_y/dt and dP_z/dt equations similar to (9.12).

We get

$$\frac{dP_x}{dt} = -\frac{e}{c}\left(\dot{y}\mathscr{H}_z - \dot{z}\mathscr{H}_y\right) - e\mathscr{E}_x = F_x$$

$$\frac{dP_y}{dt} = -\frac{e}{c}\left(\dot{z}\mathscr{H}_x - \dot{x}\mathscr{H}_z\right) - e\mathscr{E}_y = F_y$$

$$\frac{dP_z}{dt} = -\frac{e}{c}\left(\dot{x}\mathscr{H}_y - \dot{y}\mathscr{H}_x\right) - e\mathscr{E}_z = F_z \tag{9.13}$$

These equations coincide with the classical equations (2.2).

10. Angular momentum and the spin vector in Dirac's theory

Let us consider the time derivatives of the operators σ_x, σ_y, σ_z, which enter into the generalized Pauli operator. For this it is convenient to express, by means of (4.5) and (4.10), the matrices α_k in terms of ρ and σ. We can then write the Hamiltonian (9.3) as

$$H = c\rho_a\left(\sigma_x P_x + \sigma_y P_y + \sigma_z P_z\right) + mc^2\rho_c - e\Phi \tag{10.1}$$

Recalling that the σ's commute with the ρ's, we find the time derivative by means of the general formula (9.9):

$$\frac{d\sigma_x}{dt} = \frac{ic\rho_a}{\hbar}\left[\left(\sigma_y\sigma_x - \sigma_x\sigma_y\right)P_y + \left(\sigma_z\sigma_x - \sigma_x\sigma_z\right)P_z\right]$$

Whence, if we use the properties (4.6) of the σ's, we get

$$\frac{d\sigma_x}{dt} = \frac{2c}{\hbar}\rho_a\left(\sigma_z P_y - \sigma_y P_z\right) \tag{10.2}$$

Finally, if we write this formula together with the other two similar formulas, we get

$$\frac{\hbar}{2}\frac{d\sigma_x}{dt} = -\dot{y}P_z + \dot{z}P_y$$

$$\frac{\hbar}{2}\frac{d\sigma_y}{dt} = -\dot{z}P_x + \dot{x}P_z$$

$$\frac{\hbar}{2}\frac{d\sigma_z}{dt} = -\dot{x}P_y + \dot{y}P_x \tag{10.3}$$

where we have used the equations of motion (9.10) and also the relationships (4.10), which express the α_k in terms of ρ and σ.

In classical mechanics P_x, P_y, P_z are proportional to $\dot{x}$, $\dot{y}$, $\dot{z}$, which means the right-hand sides in (10.3) would vanish. The left-hand sides would also vanish because the transition to classical mechanics is equivalent to $\hbar$ being zero. We note that the order of the multipliers in (10.3) is irrelevant.

To return to quantum mechanics. We write the right-hand side of the first equation in (10.3) as

$$- \dot{y}P_z + \dot{z}P_y = - \frac{d}{dt}(yP_z - zP_y) + y\frac{dP_z}{dt} - z\frac{dP_y}{dt}$$

If we now substitute the expressions for $\dot{P}_y$ and $\dot{P}_z$ by means of Eqs. (9.13), we find that Eqs. (10,3) yield

$$\frac{d}{dt}\left(yP_z - zP_y + \frac{\hbar}{2}\,\sigma_x\right) = yF_z - zF_y \qquad (10.4)$$

and similarly

$$\frac{d}{dt}\left(zP_x - xP_z + \frac{\hbar}{2}\,\sigma_y\right) = zF_x - xF_z$$

$$\frac{d}{dt}\left(xP_y - yP_x + \frac{\hbar}{2}\sigma_z\right) = xF_y - yF_x \qquad (10.4^*)$$

We can interpret these equations as the counterpart of the law of classical mechanics that states that the time derivative of angular momentum is the torque, or the moment of force. The angular momentum in our problem is

$$\mathcal{M}_x = yP_z - zP_y + \frac{\hbar}{2}\,\sigma_x$$

$$\mathcal{M}_y = zP_x - xP_z + \frac{\hbar}{2}\,\sigma_y$$

$$\mathcal{M}_z = xP_y - yP_x + \frac{\hbar}{2}\,\sigma_z \qquad (10.5)$$

This is a generalization of the expressions that we studied in detail in the part devoted to Pauli's theory. The relations (10.5) transform into those of Pauli's theory if we nullify the vector potential (so that $P_x = p_x$, $P_y = p_y$, $P_z = p_z$) and represent the operators σ_x, σ_y, σ_z in the form of the 2×2 Pauli matrices.

Let us now try to generalize the operator

$$P = \sigma_x P_x + \sigma_y P_y + \sigma_z P_z \qquad (10.6)$$

(We studied this operator in Pauli's theory.) We construct the time derivative of P. If we turn to the Hamiltonian (10.1), we see that we can express it in terms of P as

$$H = c\rho_a P + mc^2\rho_c - e\Phi \qquad (10.7)$$

This implies that the only term in H that does not commute with P is the one with the scalar potential. Whence

$$\frac{dP}{dt} = \frac{\partial P}{\partial t} + \frac{ie}{\hbar}(P\Phi - \Phi P) = \frac{e}{c}\left(\sigma_x\frac{\partial A_x}{\partial t} + \sigma_y\frac{\partial A_y}{\partial t} + \sigma_z\frac{\partial A_z}{\partial t}\right)$$

$$+ e\left(\sigma_x\frac{\partial \Phi}{\partial x} + \sigma_y\frac{\partial \Phi}{\partial y} + \sigma_z\frac{\partial \Phi}{\partial z}\right)$$

or

$$\frac{dP}{dt} = - e\left(\sigma_x \mathscr{E}_x + \sigma_y \mathscr{E}_y + \sigma_z \mathscr{E}_z\right) \tag{10.8}$$

Thus the time derivative of P is proportional to the scalar product of the electric field $\mathscr{E}$ and the spin vector σ. When the electric field is zero, P is a constant of the motion. Note that in Pauli's theory we would have obtained the same equation of motion for P, but there the Hamiltonian depended on P quadratically and not linearly.

In Pauli's theory we also dealt with the operator

$$\mathscr{M} = \sigma_x m_x + \sigma_y m_y + \sigma_z m_z + \hbar \tag{10.9}$$

We likewise showed that this operator anticommutes with P. For this reason in Dirac's theory $\mathscr{M}$ will not be a constant of the motion even for a free electron. But since matrix ρ_c anticommutes with matrix ρ_a, which stands in the first member of the Hamiltonian (10.7), and commutes with the other two members of (10.7) the operator $\mathscr{M}_D = \rho_c \mathscr{M}$ will, in the absence of a field, commute with all the members of (10.7) and thus will be a constant of the motion.

When there is no electromagnetic field, $\mathscr{M}$ can be written as

$$\mathscr{M} = \sigma_x \mathscr{M}_x + \sigma_y \mathscr{M}_y + \sigma_z \mathscr{M}_z - \frac{\hbar}{2} \tag{10.10}$$

where $\mathscr{M}_x$, $\mathscr{M}_y$, $\mathscr{M}_z$ are determined by (10.5). We will assume that $\mathscr{M}$ is determined by (10.10) also when the momentum operators P_x, P_y, P_z contain the vector potential. Let us construct the expression for the total time derivative of operator $\mathscr{M}_D = \rho_c \mathscr{M}$ for this (general) case. We will have

$$\frac{d\mathscr{M}_D}{dt} = \frac{d}{dt}\left(\rho_c \mathscr{M}\right)$$
$$= - e\rho_c\left[\sigma_x\left(y\mathscr{E}_z - z\mathscr{E}_y\right) + \sigma_y\left(z\mathscr{E}_x - x\mathscr{E}_z\right) + \sigma_z\left(x\mathscr{E}_y - y\mathscr{E}_x\right)\right]$$
$$+ e\rho_b\left[\sigma_x\left(y\mathscr{H}_z - z\mathscr{H}_y\right) + \sigma_y\left(z\mathscr{H}_x - x\mathscr{H}_z\right) + \sigma_z\left(x\mathscr{H}_y - y\mathscr{H}_x\right)\right] \tag{10.11}$$

The right-hand side vanishes not only in the absence of a field but also when the magnetic field is zero and the electric field is directed along the radius (a central field), which is an important case for applications. We will consider the problem of an electron in a central field in the next chapter.

11. The kinetic energy of an electron

If in (9.3) or (10.1) we drop the term with the scalar potential, we come to the operator

$$T = c\,(\alpha_1 P_x + \alpha_2 P_y + \alpha_3 P_z) + mc^2\alpha_4 \tag{11.1}$$

or

$$T = c\rho_a\,(\sigma_x P_x + \sigma_y P_y + \sigma_z P_z) + mc^2\rho_c \tag{11.2}$$

which can be interpreted as the kinetic-energy operator in contrast to the Hamiltonian H, which is the total-energy operator. This new operator is the quantum counterpart of the classical quantity

$$T = \frac{mc^2}{(1 - v^2/c^2)^{1/2}} = c\,(m^2c^2 + \mathbf{p}^2)^{1/2} \tag{11.3}$$

where v is the speed of the electron, and $\mathbf{p}$ is its momentum. Such an interpretation is justified by (a) that there is an analogy between the quantum and classical expressions for the time derivative of T and (b) that the eigenvalues of T are greater than mc^2 in absolute value. This analogy becomes still greater if instead of operator T we associate with the classical kinetic energy the mean value of the Heisenberg matrix for this operator. We will find the mean value of the matrix in the next section.

Let us first construct the time derivative of operator T. By the general formula we have

$$\frac{dT}{dt} = \frac{\partial T}{\partial t} + \frac{i}{\hbar}\,(HT - TH) \tag{11.4}$$

where

$$H = T - e\Phi \tag{11.5}$$

But T depends on time only through the vector potential, which appears in the operators P_x, P_y, P_z. On the other hand, the only term in H that does not commute with T is $-e\Phi$. For this reason

$$\frac{dT}{dt} = e\left(\alpha_1 \frac{\partial A_x}{\partial t} + \alpha_2 \frac{\partial A_y}{\partial t} + \alpha_3 \frac{\partial A_z}{\partial t}\right) + ec\left(\alpha_1 \frac{\partial \Phi}{\partial x} + \alpha_2 \frac{\partial \Phi}{\partial y} + \alpha_3 \frac{\partial \Phi}{\partial z}\right) \tag{11.6}$$

or

$$\frac{dT}{dt} = -\,ec\,(\alpha_1\mathscr{E}_x + \alpha_2\mathscr{E}_y + \alpha_3\mathscr{E}_z) \tag{11.6*}$$

But in Section 9 we have seen that according to the Dirac equation

$$\dot{x} = c\alpha_1, \qquad \dot{y} = c\alpha_2, \qquad \dot{z} = c\alpha_3 \tag{11.7}$$

Whence we can write (11.6*) as

$$\frac{dT}{dt} = - e \left(\dot{x} \mathscr{E}_x + \dot{y} \mathscr{E}_y + \dot{z} \mathscr{E}_z \right) \tag{11.8}$$

This equation formally coincides with Eq. (2.3)

To make sure that the eigenvalues of T exceed mc^2 in absolute value we must find the square of T. If we use (11.2) and the properties of the matrices ρ, we get

$$T^2 = m^2 c^4 + c^2 P^2 \tag{11.9}$$

The second term on the right-hand side is c^2 times the square of the hermitian operator

$$P = \sigma_x P_x + \sigma_y P_y + \sigma_z P_z \tag{11.10}$$

[we studied this operator in the part devoted to Pauli's theory (Part III)]. If we denote the eigenvalues (real quantities) by P', the eigenvalues of T^2 will be

$$T'^2 = c^2 \left(m^2 c^2 + P'^2 \right) \tag{11.11}$$

so that

$$T' = \pm c \left(m^2 c^2 + P'^2 \right)^{1/2} \tag{11.12}$$

and hence

$$\lvert T' \rvert \geqslant mc^2 \tag{11.13}$$

Let us show that the theory does give two signs of the eigenvalues of T. We write the eigenvalue equation

$$T\psi = c\rho_a \left(\sigma_x P_x + \sigma_y P_y + \sigma_z P_z \right) \psi + mc^2 \rho_c \psi = T'\psi \tag{11.14}$$

If ψ is a solution of this equation corresponding to eigenvalue T', then

$$\psi^* = \rho_b \psi \tag{11.15}$$

will be the eigenfunction corresponding to eigenvalue $-T'$. Indeed, in virtue of the fact that matrix ρ_b commutes with matrices σ_x, σ_y, σ_z and anticommutes with ρ_a and ρ_c we find that

$$T\psi^* = T\rho_b \psi = - \rho_b T\psi = - \rho_b T'\psi = - T'\rho_b \psi$$

that is,

$$T\psi^* = - T'\psi^* \tag{11.16}$$

which proves our statement.

12. The second intrinsic degree of freedom of the electron

The fact that the theory predicts negative values for kinetic energy presents a substantial difficulty. This difficulty is connected with the above-mentioned paradox that the eigenvalues of the velocity operators for an electron are $\pm c$. Both corollaries are due

to a second degree of freedom of the electron, and this is described by the operators ρ_a, ρ_b, ρ_c (we consider the first degree of freedom to be described by the spin vector with components σ_x, σ_y, σ_z). This second degree of freedom is of a relativistic nature. Apparently, its physical meaning is that the Dirac equation, in a sense, describes not only electrons but positrons as well (a *positron* is an elementary particle with electron mass and positive charge equal to that of the electron).

If we are to interpret the Dirac equation in such a way, it is impossible to preserve the usual interpretation of the wave function as describing a one-particle state. The usual interpretation can still be applied, in a sense, to quantities that are not connected with transitions from states with positive energies to states with negative energies (or with the inverse transitions). (This concerns first of all the corresponding elements of Heisenberg's matrices.) If we want to isolate such quantities, we must build for each operator the corresponding Heisenberg matrix and then discard the matrix elements that correspond to transitions between states with energies of opposite signs (that is, between energies of the order of $+mc^2$ and $-mc^2$). Since these elements contain rapidly oscillating factors $e^{\pm i\omega t}$, where ω is of the order of $2mc^2/\hbar$ discarding these matrix elements is equivalent to averaging over a time interval that is much greater than $1/\omega$ but smaller than the reciprocal of the frequencies of ordinary transitions. Let us illustrate this by the following example.

We construct Heisenberg's matrices of the operators ρ_a, ρ_b, ρ_c, which enter into the Hamiltonian

$$H = c\rho_a \left(\sigma_x P_x + \sigma_y P_y + \sigma_z P_z \right) + mc^2\rho_c - e\Phi \qquad (12.1)$$

According to (4.8) these operators satisfy the following relationships:

$$\rho_a^2 = 1, \qquad \rho_b^2 = 1, \qquad \rho_c^2 = 1$$

$$\rho_b\rho_c = i\rho_a, \qquad \rho_c\rho_a = i\rho_b, \qquad \rho_a\rho_b = i\rho_c \qquad (12.2)$$

If we then use the notation

$$P = \sigma_x P_x + \sigma_y P_y + \sigma_z P_z \qquad (12.3)$$

we can write the Hamiltonian as

$$H = c\rho_a P + mc^2\rho_c - e\Phi \qquad (12.4)$$

By using formula (10.8), that is,

$$\frac{dP}{dt} = -e \left(\sigma_x \mathscr{E}_x + \sigma_y \mathscr{E}_y + \sigma_z \mathscr{E}_z \right) \qquad (12.5)$$

we can maintain that, in the absence of an electric field, P is a constant of the motion.

In order to build Heisenberg's matrices for the operators ρ_a, ρ_b, ρ_c we construct the equations of motion for them. If for the sake of brevity we put

$$\omega = \frac{2mc^2}{\hbar} \tag{12.6}$$

we get

$$\frac{d\rho_a}{dt} = -\omega\rho_b, \qquad \frac{d\rho_b}{dt} = \omega\rho_a - \omega\frac{P}{mc}\rho_c, \qquad \frac{d\rho_c}{dt} = \omega\frac{P}{mc}\rho_b \tag{12.7}$$

The rest of this section is devoted to the case without an electric field. Since here P is a constant of the motion, in Eqs. (12.7) we can consider P to be this constant and not the operator (12.3). Then the equations will have constant coefficients and can be solved in a straightforward manner.

Instead of ρ_a, ρ_b, ρ_c we introduce three matrices

$$\tau_a = \frac{1}{[1 + P^2/(m^2c^2)]^{1/2}}\left(\rho_a - \frac{P}{mc}\rho_c\right)$$

$$\tau_b = \rho_b$$

$$\tau_c = \frac{1}{[1 + P^2/(m^2c^2)]^{1/2}}\left(\frac{P}{mc}\rho_a + \rho_c\right) \tag{12.8}$$

These matrices satisfy the same relationships (12.2) as ρ_a, ρ_b, and ρ_c, namely

$$\tau_a^2 = 1, \qquad \tau_b^2 = 1, \qquad \tau_c^2 = 1$$

$$\tau_b\tau_c = i\tau_a, \qquad \tau_c\tau_a = i\tau_b, \qquad \tau_a\tau_b = i\tau_c \tag{12.9}$$

We can express the ρ's in terms of the τ's in the following way:

$$\rho_a = \frac{1}{[1 + P^2/(m^2c^2)]^{1/2}}\left(\tau_a + \frac{P}{mc}\tau_c\right)$$

$$\rho_b = \tau_b$$

$$\rho_c = \frac{1}{[1 + P^2/(m^2c^2)]^{1/2}}\left(-\frac{P}{mc}\tau_a + \tau_c\right) \tag{12.10}$$

where the constant factors differ from the ones in (12.8) only in the sign of P.

The equations of motion for the τ's are

$$\frac{d\tau_a}{dt} = -\nu\tau_b, \qquad \frac{d\tau_b}{dt} = \nu\tau_a, \qquad \frac{d\tau_c}{dt} = 0 \tag{12.11}$$

where for the sake of brevity we have put

$$\nu = \omega\left(1 + \frac{P^2}{m^2c^2}\right)^{1/2} = \frac{2c}{\hbar}(m^2c^2 + P^2)^{1/2} \tag{12.12}$$

20*

We can now easily solve Eqs. (12.11) if we take into account conditions (12.9). We have

$$\tau_a = -\tau_1 \sin \nu t - \tau_2 \cos \nu t$$
$$\tau_b = \tau_1 \cos \nu t - \tau_2 \sin \nu t$$
$$\tau_c = \tau_3 \qquad\qquad (12.13)$$

where τ_1, τ_2, τ_3 are constant matrices that satisfy the conditions

$$\tau_1^2 = 1, \qquad \tau_2^2 = 1, \qquad \tau_3^2 = 1$$
$$\tau_2\tau_3 = i\tau_1, \qquad \tau_3\tau_1 = i\tau_2, \qquad \tau_1\tau_2 = i\tau_3 \qquad (12.14)$$

similar to (12.9).

Let us apply the averaging process to the Heisenberg matrices just obtained. (We mentioned the need for this process at the beginning of this section.) We have

$$\langle \tau_a \rangle = 0, \qquad \langle \tau_b \rangle = 0, \qquad \langle \tau_c \rangle = \tau_3 \qquad (12.15)$$

Thus the mean values of Heisenberg's matrices for the operators ρ are

$$\langle \rho_a \rangle = \frac{P}{(m^2c^2 + P^2)^{1/2}} \tau_3, \qquad \langle \rho_b \rangle = 0, \qquad \langle \rho_c \rangle = \frac{mc}{(m^2c^2 + P^2)^{1/2}} \tau_3 \qquad (12.16)$$

Substituting these values into the expression for the Hamiltonian (12.4), we get

$$H = c\tau_3 (m^2c^2 + P^2)^{1/2} - e\Phi \qquad (12.17)$$

If we remember that τ_3 is a constant whose square is unity, the last formula is the same as that of the classical kinetic energy, (11.3).

13. Second-order equations

The Dirac equation is a system of four differential equations of the first order in four unknown functions. We can exclude two functions and construct a system of two second-order equations in two unknown functions. If we then perform a limiting process by tending c to infinity, we come to a nonrelativistic wave equation for an electron in a magnetic field. This brings us to Pauli's theory, which we considered in Part III.

Since the derivation of the Pauli equation from the Dirac equation is interesting in itself, we will do it here although we already know the result.

In order to find a system of two second-order equations using the Dirac equation we write the latter as

$$T\psi = e\Phi\psi + i\hbar \frac{\partial \psi}{\partial t} \qquad (13.1)$$

where T is the electron's kinetic-energy operator considered in Section 11. Let us apply operator T to both sides of this equation. After some manipulations we get

$$T^2\psi = \left[-i\hbar \frac{\partial T}{\partial t} + e(T\Phi - \Phi T) \right] \psi + i\hbar e \frac{\partial \Phi}{\partial t} \psi$$
$$+ e^2\Phi^2\psi + 2i\hbar e\Phi \frac{\partial \psi}{\partial t} - \hbar^2 \frac{\partial^2 \psi}{\partial t^2} \quad (13.2)$$

The expression in brackets on the right-hand side is, as can be seen, $-i\hbar$ times the total time derivative of operator T, which we have already calculated [see (11.5)]. We also calculated the expression for T^2, but we must modify it. If we determine P^2 in the second term (11.9), we get on the basis of (9.11) and the properties of matrices σ the following:

$$T^2 = m^2c^4 + c^2\left(P_x^2 + P_y^2 + P_z^2\right) + \hbar ec\left(\sigma_x \mathcal{H}_x + \sigma_y \mathcal{H}_y + \sigma_z \mathcal{H}_z\right) \quad (13.3)$$

We use the expressions of type

$$P_x^2 = \left(p_x + \frac{e}{c} A_x\right)^2 = p_x^2 + \frac{2e}{c} A_x p_x - \frac{i\hbar e}{c} \frac{\partial A_x}{\partial x} + \frac{e^2}{c^2} A_x^2$$

and also for brevity the vector notations, and we get

$$T^2 = m^2c^4 + c^2 p^2 + 2ec(\mathbf{A} \cdot \mathbf{p}) - i\hbar ec \operatorname{div} \mathbf{A} + e^2\mathbf{A}^2 + \hbar ec(\sigma \cdot \mathcal{H}) \quad (13.4)$$

We insert this expression into (13.2) and use (11.8) and the Lorentz condition

$$\operatorname{div} \mathbf{A} + \frac{1}{c} \frac{\partial \Phi}{\partial t} = 0 \quad (13.5)$$

This yields

$$[m^2c^4 + c^2 p^2 + 2ec(\mathbf{A} \cdot \mathbf{p}) + e^2(\mathbf{A}^2 - \Phi^2) + \hbar ec(\sigma \cdot \mathcal{H}) + i\hbar e(\mathbf{x} \cdot \mathcal{E})] \psi$$
$$= 2i\hbar e\, \Phi \frac{\partial \psi}{\partial t} - \hbar^2 \frac{\partial^2 \psi}{\partial t^2} \quad (13.6)$$

This equation can also be written as

$$-\nabla^2\psi + \frac{1}{c^2} \frac{\partial^2 \psi}{\partial t^2} - \frac{2ie}{\hbar c}\left[(\mathbf{A} \cdot \operatorname{grad}\psi) + \frac{\Phi}{c} \frac{\partial \psi}{\partial t} \right] + \frac{m^2c^2}{\hbar^2} \psi$$
$$+ \frac{e^2}{\hbar^2 c^2}(\mathbf{A}^2 - \Phi^2)\psi + \frac{e}{\hbar c}(\sigma \cdot \mathcal{H})\psi + \frac{ie}{\hbar c^2}(\dot{\mathbf{x}} \cdot \mathcal{E})\psi = 0 \quad (13.7)$$

This is a system of four equations for the four functions ψ_1, ψ_2, ψ_3, ψ_4. With our choice of matrices α_k the first two equations include only the first two functions ψ_1 and ψ_2, and the second two equations only ψ_3 and ψ_4, so that Eq. (13.7) separates into two systems.

Equation (13.7) differs from the relativistic generalization of the Schrödinger equation proposed by various authors before the concept of the electron spin was introduced and before Dirac's

theory. The difference lies in the two last terms, which contain the matrices σ and $\dot{x}$.

Let us see what equation we get from (13.6) or (13.7) if we omit the relativistic correction, that is, if we effect the limiting process $c \to \infty$ on the assumption that the energy of the particle is close to its rest energy mc^2.

For this we put

$$\psi = \psi^* e^{-imc^2t/\hbar} \tag{13.8}$$

and assume that ψ^* varies slowly in time in comparison with ψ.

For stationary states this assumption means that in

$$\psi = \psi^0 e^{-iWt/\hbar} \tag{13.9}$$

we have put

$$W = mc^2 + E \tag{13.10}$$

and consider E to be very small in comparison with mc^2. If we insert (13.8) into Eq. (13.6) and divide the result by $2mc^2$, we get the exact equation

$$\left[\frac{1}{2m}\left(\mathbf{p} + \frac{e}{c}\mathbf{A}\right)^2 - e\Phi + \frac{\hbar e}{2mc}(\sigma \cdot \mathscr{H})\right]\psi^* - i\hbar\frac{\partial\psi^*}{\partial t}$$
$$= \frac{1}{2mc^2}\left[-i\hbar e(\dot{\mathbf{x}}\cdot\mathscr{E}) + \left(e\Phi + i\hbar\frac{\partial}{\partial t}\right)^2\right]\psi^* \tag{13.11}$$

We must remember that when passing to the limit $c \to \infty$ the factor $1/c$ in the terms containing the magnetic quantities $\mathbf{A}$ and $\mathscr{H}$ appears from the use of the Gaussian units, that is, the factor is a constant. Because of this we must retain all the terms on the left-hand side of (13.11), whereas we replace the right-hand side with zero. The approximate equation is then

$$\left[\frac{1}{2m}\left(\mathbf{p} + \frac{e}{c}\mathbf{A}\right)^2 - e\Phi + \frac{\hbar e}{2mc}(\sigma \cdot \mathscr{H})\right]\psi^* = i\hbar\frac{\partial\psi^*}{\partial t} \tag{13.12}$$

The operator on the left-hand side of (13.12)

$$H^* = \frac{1}{2m}\left(\mathbf{p} + \frac{e}{c}\mathbf{A}\right)^2 - e\Phi + \frac{\hbar e}{2mc}(\sigma \cdot \mathscr{H}) \tag{13.13}$$

is hermitian. If we take σ to be the set of the 2×2 Pauli matrices, the Hamiltonian H^* will coincide with the Pauli operator examined in Section 5, Part III, and the wave equation (13.12) will coincide with the Pauli equation. Introduction of 4×4 matrices makes no essential change because the equation for four component functions, (13.12), splits into two equivalent systems of equations for two component functions.

The fact that the Pauli equation is obtained from the Dirac equation as an approximation is additional proof that it is valid.

In conclusion we write the wave equation (13.12) in explicit form. If by H^0 in accordance with (5.18) Part III, we denote

$$H^0 = \frac{1}{2m}\left(\mathbf{p} + \frac{e}{c}\,\mathbf{A}\right)^2 - e\Phi \tag{13.14}$$

which does not contain matrices, we get

$$H^* = H^0 + \mu^0\,(\sigma \cdot \mathscr{H}) \tag{13.15}$$

where

$$\mu^0 = \frac{e\hbar}{2mc} \tag{13.16}$$

is the Bohr magneton.

If we assume that σ are the 2×2 Pauli matrices introduced in Section 1, Part III,

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{13.17}$$

and ψ^* is the two component wave function (the first two components of the four component Dirac function), we can with our choice of matrices $\sigma_x, \sigma_y, \sigma_z$ [see (4.26)] write Eq. (13.12) as

$$H^0\psi_1^* + \mu^0\left[\mathscr{H}_z\psi_1^* + \left(\mathscr{H}_x - i\mathscr{H}_y\right)\psi_2^*\right] - i\hbar\,\frac{\partial\psi_1^*}{\partial t} = 0$$

$$H^0\psi_2^* + \mu^0\left[\left(\mathscr{H}_x + i\mathscr{H}_y\right)\psi_1^* - \mathscr{H}_z\psi_2^*\right] - i\hbar\,\frac{\partial\psi_2^*}{\partial t} = 0 \tag{13.18}$$

The equations for the last two components (ψ_3^* and ψ_4^*) will be similar. They will differ from (13.18) in the signs of the terms that derive from σ_1 and σ_3 (in other words, the terms proportional to $\mathscr{H}_x$ and $\mathscr{H}_z$) If we choose the matrices according to Dirac [see (4.18)], the equations for the last two components will be a simple repetition of the equations for the first two.

Chapter II

THE USE OF THE DIRAC EQUATION
IN PHYSICAL PROBLEMS

1. The free electron

The wave equation for a free electron has the form

$$H\psi - i\hbar \frac{\partial \psi}{\partial t} = 0 \tag{1.1}$$

where according to (3.7) and (4.13), Chapter I,

$$H = c\left(\alpha_1 p_x + \alpha_2 p_y + \alpha_3 p_z\right) + mc^2 \alpha_4 \tag{1.2}$$

or

$$H = c\rho_a\left(\sigma_x p_x + \sigma_y p_y + \sigma_z p_z\right) + mc^2 \rho_c \tag{1.3}$$

Since the law of conservation of energy holds for a free electron, we can add to the wave equation the eigenvalue equation for the Hamiltonian

$$H\psi = W\psi \tag{1.4}$$

Further, the operators p_x, p_y, p_z commute with H and therefore are constants of the motion. Since they also commute with each other, we can consider the momentum components to be given numbers p'_x, p'_y, p'_z and subject ψ to the additional conditions

$$p_x \psi = -i\hbar \frac{\partial \psi}{\partial x} = p'_x \psi$$

$$p_y \psi = -i\hbar \frac{\partial \psi}{\partial y} = p'_y \psi$$

$$p_z \psi = -i\hbar \frac{\partial \psi}{\partial z} = p'_z \psi \tag{1.5}$$

Mathematically this means we have assumed the dependence of all four functions ψ_i on the coordinates and time to be

$$\psi = \psi^0 \exp\left[\frac{i}{\hbar}\left(xp'_x + yp'_y + zp'_z - Wt\right)\right] \tag{1.6}$$

that is, we are considering a plane wave.

Another constant of the motion is the operator

$$P = \sigma_x p_x + \sigma_y p_y + \sigma_z p_z \tag{1.7}$$

which commutes with both H and p_x, p_y, p_z. Hence we can subject ψ to yet another condition

$$P\psi = P'\psi \tag{1.8}$$

We met operator P in Pauli's theory, but there we had no occasion to calculate its eigenfunctions since the Pauli equation contains only its square, which in the absence of a field is

$$P^2 = p_x^2 + p_y^2 + p_z^2 \tag{1.9}$$

For this reason, when p_x', p_y', p_z' are given, eigenvalue P' can only have two values:

$$P' = + (p_x'^2 + p_y'^2 + p_z'^2)^{1/2} \quad \text{and} \quad P' = - (p_x'^2 + p_y'^2 + p_z'^2)^{1/2} \tag{1.10}$$

The Hamiltonian expressed in terms of P will have the form

$$H = c\rho_a P + mc^2 \rho_c \tag{1.11}$$

Since

$$H^2 = m^2 c^4 + c^2 P^2 \tag{1.12}$$

the eigenvalues W of H will be

$$W = + (m^2 c^4 + c^2 P^2)^{1/2} \quad \text{and} \quad W = - (m^2 c^4 + c^2 P^2)^{1/2} \tag{1.13}$$

Thus for a given value of momentum we have four solutions:

$$
\begin{array}{llll}
\text{first} & W = +|W|, & P = +|P| \\
\text{second} & W = +|W|, & P = -|P| \\
\text{third} & W = -|W|, & P = +|P| \\
\text{fourth} & W = -|W|, & P = -|P| & (1.14)
\end{array}
$$

The first two correspond to positive kinetic energy. Of these the first corresponds to the magnetic moment or the spin vector coinciding in direction with the direction of motion; the second corresponds to motion in the opposite direction. The second two correspond to negative energy and have no physical meaning in ordinary quantum mechanics, which deals only with fixed number of charged particles. In the general case we established the existence of these solutions in Section 11, Chapter I.

Now let us find the eigenfunctions that describe these four states.

We have a set of simultaneous algebraic equations (1.4) and (1.8), which we write (ignoring the primes) as

$$(\sigma_x p_x + \sigma_y p_y + \sigma_z p_z) \psi = P\psi \tag{1.15}$$

$$(c\rho_a P + mc^2 \rho_c) \psi = W\psi \tag{1.16}$$

These serve to determine the four component function ψ. But Eq. (1.15) also holds for the two component function of Pauli's theory.

If according to formulas in Section 1, Chapter III, we consider σ_x, σ_y, σ_z to be the Pauli matrices

$$\sigma_x = \sigma_1, \qquad \sigma_y = \sigma_2, \qquad \sigma_z = \sigma_3 \tag{1.17}$$

we can write (1.15) as

$$(p_x - ip_y)\,\psi_2 + p_z\psi_1 = P\psi_1$$
$$(p_x + ip_y)\,\psi_1 - p_z\psi_2 = P\psi_2 \tag{1.18}$$

Owing to the relationship (1.9) the system determinant vanishes. We can put

$$\psi_1 = \lambda\,(P + p_z), \qquad \psi_2 = \lambda\,(p_x + ip_y) \tag{1.19}$$

where λ is constant.

If, on the other hand, we view (1.15) as the equation for a four component function and, in accordance with formulas (4.26), Chapter I, take for σ_x, σ_y, σ_z the matrices

$$\sigma_x = s_1, \qquad \sigma_y = s_2, \qquad \sigma_z = s_3 \tag{1.20}$$

where

$$s_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1, & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{pmatrix}, \quad s_2 = \begin{pmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix}, \quad s_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\tag{1.21}$$

then the equations for the functions ψ_1 and ψ_2, (1.18), will keep their form but will be joined by two similar equations for functions ψ_3 and ψ_4, namely

$$-(p_x + ip_y)\,\psi_4 - p_z\psi_3 = P\psi_3$$
$$-(p_x - ip_y)\,\psi_3 + p_z\psi_4 = P\psi_4 \tag{1.22}$$

We can write the solution to these equations as

$$\psi_3 = \mu\,(p_x + ip_y), \qquad \psi_4 = -\mu\,(P + p_z) \tag{1.23}$$

Hence the solution of (1.15) for the four component function ψ will contain two arbitrary constants, λ and μ. Their ratio can be determined from (1.16).

With our choice of matrices we have, according to (4.27), Chapter I, the following:

$$\rho_a = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \quad \rho_b = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix}, \quad \rho_c = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}$$

$$\tag{1.24}$$

and Eq. (1.16) in expanded form will be

$$cP\psi_1 - mc^2\psi_4 = W\psi_1$$
$$cP\psi_2 + mc^2\psi_3 = W\psi_2$$
$$-cP\psi_3 + mc^2\psi_2 = W\psi_3$$
$$-cP\psi_4 - mc^2\psi_1 = W\psi_4 \tag{1.25}$$

Expressing by (1.19) and (1.23) the components of the wave function in terms of λ and μ, we get two equations

$$(W - cP)\,\lambda - mc^2\mu = 0$$

$$-mc^2\lambda + (W + cP)\,\mu = 0 \qquad (1.26)$$

each of which appears two times in (1.25). These equations yield

$$\frac{\lambda}{\mu} = \frac{W + cP}{mc^2} = \frac{mc^2}{W - cP} \qquad (1.27)$$

It follows that the ratio λ/μ is real and its sign coincides with the sign of W. Another corollary of (1.27) is

$$\frac{\lambda^2 + \mu^2}{2\lambda\mu} = \frac{W}{mc^2}, \qquad \frac{\lambda^2 - \mu^2}{2\lambda\mu} = \frac{P}{mc} \qquad (1.28)$$

Substituting the found values (1.19) and (1.23) of the components of the wave function into the expressions for the current density given in Section 8, Chapter I, and using the relationship (1.28), we get

$$A_4 = 4\lambda\mu P\,(P + p_z), \qquad A_5 = 0 \qquad (1.29)$$

and for the space-time components of current density

$$A_0 = \frac{W}{mc^2}\,A_4, \quad A_1 = \frac{p_x}{mc}\,A_4, \quad A_2 = \frac{p_y}{mc}\,A_4, \quad A_3 = \frac{p_z}{mc}\,A_4 \,(1.30)$$

At $W > 0$ we can normalize function ψ in a way such that $A_4 = mc^2$, and at $W < 0$ so that $A_4 = -mc^2$. Then at $W > 0$

$$A_0 = W, \quad A_1 = cp_x, \quad A_2 = cp_y, \quad A_3 = cp_z \qquad (1.31)$$

and at $W < 0$

$$A_0 = -W, \quad A_1 = -cp_x, \quad A_2 = -cp_y, \quad A_3 = -cp_z \quad (1.32)$$

Hence the spatial components of the current density are proportional to the momentum, and the ratios between them and the time component correspond to the ratio of the speed of particle to the speed of light.

In conclusion we note that in the nonrelativistic limit, when W is close to $+mc^2$, the values of the constants λ and μ are close to each other and in consequence we arrive at the approximate equalities

$$\psi_1 \approx -\psi_4, \qquad \psi_2 \approx \psi_3 \quad \text{at} \quad W > 0 \qquad (1.33)$$

But if $|p| \ll mc$ and W is negative,

$$\psi_1 \approx \psi_4, \qquad \psi_2 \approx -\psi_3 \quad \text{at} \quad W < 0 \qquad (1.34)$$

2. An electron in a homogeneous magnetic field

We examined the influence of a homogeneous magnetic field on the energy levels of an electron in the nonrelativistic approximation at the end of Chapter III, devoted to Pauli's theory. Here we will consider a simple problem, assuming that no forces other than a homogeneous magnetic field act on the electron, but we will solve the problem on the basis of Dirac's theory with no further omissions.

We assume that the field is directed along the z axis and in absolute value is equal to $|\mathcal{H}|$. We can put the vector potential equal to

$$A_x = -\frac{1}{2}|\mathcal{H}|y, \qquad A_y = \frac{1}{2}|\mathcal{H}|x, \qquad A_z = 0 \qquad (2.1)$$

which in cylindrical (polar) coordinates

$$x = \rho \cos \varphi, \qquad y = \rho \sin \varphi \qquad (2.2)$$

correspond to

$$A_\varphi = \frac{1}{2}|\mathcal{H}|\rho^2, \qquad A_\rho = 0, \qquad A_z = 0 \qquad (2.1^*)$$

We will formulate our problem in Cartesian coordinates and only at the end of the calculations shift to polar (non-Cartesian) coordinates.

The Hamiltonian will have the form

$$H = c\rho_a (\sigma_x P_x + \sigma_y P_y + \sigma_z P_z) + mc^2\rho_c \qquad (2.3)$$

where

$$P_x = p_x - \frac{e}{2c}|\mathcal{H}|y, \qquad P_y = p_y + \frac{e}{2c}|\mathcal{H}|x, \qquad P_z = p_z \quad (2.4)$$

Here, as with the free electron, the operator

$$P = \sigma_x P_x + \sigma_y P_y + \sigma_z P_z \qquad (2.5)$$

which enters into the expression for H will be a constant of the motion. Another constant of the motion will be operator p_z, which commutes with P.

We can therefore consider the complete system of equations

$$H\psi = W\psi \qquad (2.6)$$

$$P\psi = P'\psi \qquad (2.7)$$

$$p_z\psi = p_z'\psi \qquad (2.8)$$

As in the case of the free electron every eigenvalue of P will have corresponding to it two values of W:

$$W = \pm (m^2c^4 + c^2P'^2)^{1/2} \qquad (2.9)$$

The problem reduces to finding the eigenfunctions of P, that is, to solving Eq. (2.7). With our choice of matrices Eq. (2.7) can be written as

$$(\sigma_1 P_x + \sigma_2 P_y + \sigma_3 P_z)\,\psi = P'\psi \qquad (2.10)$$

where σ_k are the 2×2 Pauli matrices. Once we find ψ_1 and ψ_2 from these equations, we can obtain ψ_3 and ψ_4 from (1.16).

If we apply P to (2.10) and make use of the commutation relations between P_x, P_y, P_z [see Section 5, Part III, and Section 9, Chapter I, Part VI], we get

$$\left(P_x^2 + P_y^2 + P_z^2 + \frac{\hbar e}{c}\,\sigma_3\,|\,\mathcal{H}\,|\right)\psi = P'^2\psi \qquad (2.11)$$

If we divide this equation by $2m$, we get on the left-hand side an operator that enters into Pauli's operator [see (5.19), Part III] and coincides with it when the electric field is zero.

Now Eq. (2.11) has only one matrix, σ_3, which moreover is diagonal. For this reason (2.11) splits into two equations, each of which has only one function ψ_i. Bearing in mind (2.3) and (2.8), we can write them as

$$\left[p_x^2 + p_y^2 + \frac{e\,|\,\mathcal{H}\,|}{c}\,(xp_y - yp_x + \hbar) + \frac{e^2\,|\,\mathcal{H}\,|^2}{4c^2}\,(x^2 + y^2)\right]\psi_1$$
$$= (P'^2 - p_z'^2)\,\psi_1 \qquad (2.12)$$

$$\left[p_x^2 + p_y^2 + \frac{e\,|\,\mathcal{H}\,|}{c}\,(xp_y - yp_x - \hbar) + \frac{e^2\,|\,\mathcal{H}\,|^2}{4c^2}\,(x^2 + y^2)\right]\psi_2$$
$$= (P' - p_z')\,\psi \qquad (2.12^*)$$

These equations differ only in the sign in front of the term containing $\hbar$. Let us express the operators p_x and p_y in terms of derivatives and put for brevity

$$\frac{e\,|\,\mathcal{H}\,|}{2\hbar c} = b, \qquad \frac{1}{4\hbar^2 b}\,(P'^2 - p_z'^2) = l \qquad (2.13)$$

We get

$$-\frac{\partial^2 \psi_1}{\partial x^2} - \frac{\partial^2 \psi_1}{\partial y^2} - 2ib\left(x\,\frac{\partial \psi_1}{\partial y} - y\,\frac{\partial \psi_1}{\partial x}\right) + b^2(x^2 + y^2)\,\psi_1$$
$$= 2b\,(2l - 1)\,\psi_1 \qquad (2.14)$$

$$-\frac{\partial^2 \psi_2}{\partial x^2} - \frac{\partial^2 \psi_2}{\partial y^2} - 2ib\left(x\,\frac{\partial \psi_2}{\partial y} - y\,\frac{\partial \psi_2}{\partial x}\right) + b^2(x^2 + y^2)\,\psi_2$$
$$= 2b\,(2l + 1)\,\psi^2 \qquad (2.14^*)$$

In polar coordinates, (2.2), these equations take the form

$$\frac{\partial^2 \psi_1}{\partial \rho^2} + \frac{1}{\rho}\,\frac{\partial \psi_1}{\partial \rho} + \frac{1}{\rho^2}\,\frac{\partial^2 \psi_1}{\partial \varphi^2} + 2ib\,\frac{\partial \psi_1}{\partial \varphi} - b^2\rho^2\psi_1 + 2b\,(2l - 1)\,\psi_1 = 0$$
$$(2.15)$$

$$\frac{\partial^2 \psi_2}{\partial \rho^2} + \frac{1}{\rho}\,\frac{\partial \psi_2}{\partial \rho} + \frac{1}{\rho^2}\,\frac{\partial^2 \psi_2}{\partial \varphi^2} + 2ib\,\frac{\partial \psi_2}{\partial \varphi} - b^2\rho^2\psi_2 + 2b\,(2l + 1)\,\psi_2 = 0$$
$$(2.15^*)$$

We could arrive at these equations more directly if we used the operator P transformed to cylindrical coordinates according to the formulas in Section 6, Part III.

Equations (2.15) and (2.15*) are easily solved by separating the variables. In this case it is sufficient to consider the equation for one of the functions ψ_1 or ψ_2, since owing to (2.7) these functions are connected in the following way:

$$\frac{\partial \psi_2}{\partial x} - i \frac{\partial \psi_2}{\partial y} + b(x - iy)\psi_2 = \frac{i}{\hbar}(P' - p_z')\psi_1 \qquad (2.16)$$

$$\frac{\partial \psi_1}{\partial x} + i \frac{\partial \psi_1}{\partial y} - b(x + iy)\psi_1 = \frac{i}{\hbar}(P' + p_z')\psi_2 \qquad (2.16^*)$$

which are a generalization of (1.18) when there is a magnetic field.

Equations (2.14) and (2 14*) are obtained from (2.16) and (2.16*) by excluding one of the functions, ψ_1 or ψ_2.

We put

$$\psi_2 = \lambda e^{-im\varphi} f \qquad (2.17)$$

where λ is a constant factor, m is an integer, and f depends only on ρ, and we introduce a new independent variable

$$\xi = b\rho^2 \qquad (2.18)$$

The equation for f that follows from (2.15*) will be

$$-\frac{d}{d\xi}\left(\xi \frac{df}{d\xi}\right) + \left(\frac{\xi}{4} + \frac{m^2}{4\xi}\right)f = \left(l + \frac{m+1}{2}\right)f \qquad (2.19)$$

This equation differs only in notation from the equation for functions connected with the generalized Laguerre polynomials, which was considered in Chapter V, Part II, devoted to Schrödinger's theory [see Eq. (3.3*), Chapter V, Part II]. We saw that the eigenvalues of the operator on the left-hand side of (2.19) are

$$l + \frac{m+1}{2} = \frac{|m|+1}{2} + p, \qquad p = 0, 1, 2, \ldots$$

so that l will be a nonnegative integer:

$$l = 0, 1, 2, \ldots \qquad (2.20)$$

With the value of l given, the number m will assume the values

$$m = -l, -l+1, \ldots, l-1, l \qquad (2.21)$$

At $m > 0$ we can take as eigenfunctions

$$f_{lm}(\xi) = e^{-\xi/2}\xi^{m/2}Q_l^m(\xi) \qquad (2.22)$$

and at $m < 0$

$$f_{lm}(\xi) = (-1)^m e^{-\xi/2}\xi^{-m/2}Q_{l+m}^{-m}(\xi) \qquad (2.22^*)$$

where Q are the generalized Laguerre polynomials, which we studied in detail in Section 4, Chapter V, Part II. With such a definition of $f_{lm}(\xi)$ we can put

$$\psi_2 = \lambda e^{-im\varphi} f_{lm}(\xi) \tag{2.23}$$

Function ψ_1 is expressed in terms of ψ_2 by (2.16), which in polar coordinates can be written as

$$\psi_1 = -\frac{i\hbar}{P'-p_z'}\frac{e^{-i\varphi}}{\rho}\left(\rho\frac{\partial\psi_2}{\partial\rho} - i\frac{\partial\psi_2}{\partial\varphi} + b\rho^2\psi_2\right) \tag{2.24}$$

Substituting (2.23) for ψ_2, we find that

$$\psi_1 = -\lambda\frac{i\hbar b^{1/2}}{P'-p_z'}e^{-i(m+1)\varphi}\frac{1}{\xi^{1/2}}\left[2\xi\frac{df_{lm}}{d\xi} + (\xi - m)f\right] \tag{2.25}$$

On the basis of the properties of the polynomials $Q_p^s(x)$

$$\frac{dQ_p^s(x)}{dx} = -pQ_{p-1}^{s+1}(x)$$

$$[(4.9), \text{ Chapter V, Part II}]$$

$$x\frac{dQ_p^s(x)}{dx} + sQ_p^s(x) = (p+s)Q_p^{s-1}(x)$$

$$[(4.12), \text{ Chapter V, Part II}]$$

it is easy to show that for both positive and negative values of m there is the relationship

$$\frac{1}{\xi^{1/2}}\left[2\xi\frac{df_{lm}}{d\xi} + (\xi - m)f_{lm}\right] = -2lf_{l-1\,m+1} \tag{2.26}$$

Substituting (2.26) into (2.25), we get

$$\psi_1 = \lambda\frac{i\hbar b^{1/2}}{P-p_z}2le^{-i(m+1)\varphi}f_{l-1\,m+1}(\xi) \tag{2.27}$$

For brevity we rewrite (2.27) and (2.23) as

$$\psi_1 = \lambda\psi_1^0, \qquad \psi_2 = \lambda\psi_2^0 \tag{2.28}$$

where ψ_1^0 and ψ_2^0 are the functions defined above.

Equations (2.10) for the two component function are the first two equations in the set of equations for four component functions

$$(s_1P_x + s_2P_y + s_3P_z)\psi = P'\psi \tag{2.29}$$

In (2.10) σ_1, σ_2, σ_3 are the 2×2 Pauli matrices, and in (2.29) s_1, s_2, s_3 are the 4×4 matrices (1.21). The structure of these

4×4 matrices is such that the first two equations in (2.29) coincide, as we have just noted, with Eqs. (2.10), and the last two are obtained from them by replacing ψ_1 with $-\psi_4$ and ψ_2 with ψ_3. Hence, if functions (2.28) satisfy the first two equations in (2.29), then along with

$$\psi_3 = \mu\psi_2^0, \qquad \psi_4 = -\psi_1^0 \tag{2.30}$$

they will satisfy all the equations in (2.29). This holds regardless of the form of operators P_x, P_y, P_z, for one, in the case of the free electron considered in Section 1, and (1.19) and (1.23) correspond to (2.28) and (2.30).

In view of this correspondence there is no need to repeat the manipulations of Section 1. We only need to remember (1.13), which connects the eigenvalues W and P':

$$W = \pm (m^2c^4 + c^2P'^2)^{1/2} \tag{2.31}$$

and the expression that follows from (2.13) for P':

$$P' = \pm (p_z'^2 + 4\hbar^2bl)^{1/2} \tag{2.32}$$

In conclusion we note that the dependence of the functions (2.23) and (2.27) on angle φ shows that they are the eigenfunctions of the operator

$$\mathcal{M}_z = xp_y - yp_x + \frac{\hbar}{2}\sigma_z = p_\varphi + \frac{\hbar}{2}\sigma_3 \tag{2.33}$$

for the eigenvalue

$$\mathcal{M}_z' = -\left(m + \frac{1}{2}\right)\hbar \tag{2.34}$$

Operator $\mathcal{M}_z$ commutes with all three operators H, P and p_z that enter into Eqs. (2.6), (2.7), and (2.8). This fact expresses the axial symmetry of the problem under consideration.

3. Constants of the motion in the problem with spherical symmetry

Let us examine the problem of describing electronic states in a field with spherical symmetry according to Dirac's theory. We explored the same problem according to Schrödinger's theory in Chapters IV and V, Part II. In addition, in Part III, devoted to Pauli's theory, we studied the properties of the angular momentum of an electron possessing spin. Now we will see the distinctions that Dirac's theory introduces; this theory explains the existence of doublets and gives a complete picture of the splitting of energy levels in a magnetic field.

It is convenient to follow the method used in classical mechanics, that is, to first consider the problem in rectangular Cartesian coordinates and only then go on to spherical coordinates.

In rectangular coordinates the Hamiltonian for our problem has the form

$$H = c\rho_a \left(\sigma_x p_x + \sigma_y p_y + \sigma_z p_z \right) + mc^2 \rho_c + U(r) \tag{3.1}$$

or

$$H = c\rho_a P + mc^2 \rho_c + U(r) \tag{3.2}$$

where

$$P = \sigma_x p_x + \sigma_y p_y + \sigma_z p_z \tag{3.3}$$

is the operator introduced when we considered the Pauli equation [see (5.11), Part III]. The difference here is only that in Dirac's theory the operators for the components of spin, σ_x, σ_y, σ_z, are represented by 4×4 matrices and in Pauli's theory by 2×2 matrices.

Pauli's theory introduces the operators

$$\mathcal{M}_x = m_x + \frac{1}{2} \hbar \sigma_x$$

$$\mathcal{M}_y = m_y + \frac{1}{2} \hbar \sigma_y$$

$$\mathcal{M}_z = m_z + \frac{1}{2} \hbar \sigma_z \tag{3.4}$$

for the components of total (orbital and spin) angular momentum. These operators satisfy the commutation relations

$$\mathcal{M}_y \mathcal{M}_z - \mathcal{M}_z \mathcal{M}_y = i\hbar \mathcal{M}_x$$

$$\mathcal{M}_z \mathcal{M}_x - \mathcal{M}_x \mathcal{M}_z = i\hbar \mathcal{M}_y$$

$$\mathcal{M}_x \mathcal{M}_y - \mathcal{M}_y \mathcal{M}_x = i\hbar \mathcal{M}_z \tag{3.5}$$

The operator composed of the three components, namely

$$\mathcal{M} = \sigma_x \mathcal{M}_x + \sigma_y \mathcal{M}_y + \sigma_z \mathcal{M}_z - \frac{1}{2} \hbar \tag{3.6}$$

can be represented as

$$\mathcal{M} = \sigma_x m_x + \sigma_y m_y + \sigma_z m_z + \hbar \tag{3.7}$$

We also see that the last operator commutes with each of the operators $\mathcal{M}_x$, $\mathcal{M}_y$, $\mathcal{M}_z$. In addition, as we saw in Section 5, Part III, the operator $\mathcal{M}$ anticommutes with P, which is determined by (3.3):

$$\mathcal{M}P + P\mathcal{M} = 0 \tag{3.8}$$

The Hamiltonian (3.2) in Dirac's theory includes operator P multiplied by matrix ρ_a, and also two terms that commute with ρ_c.

Since the matrices ρ_a and ρ_c anticommute, it follows directly that the operator

$$\mathcal{M}_D = \rho_c\mathcal{M} = \mathcal{M}\rho_c \qquad (3.9)$$

will commute with $\rho_a P$ and hence with all the terms in H, so that we get

$$H\mathcal{M}_D - \mathcal{M}_D H = 0 \qquad (3.10)$$

and consequently

$$\frac{d}{dt}(\mathcal{M}_D) = 0 \qquad (3.11)$$

Hence the quantity corresponding to operator $\mathcal{M}_D$ will be constant for a field with spherical symmetry. We calculated the time derivative of this quantity for an arbitrary field in Section 10, Chapter I [formula (10.11)].

We have seen that in the problem with spherical symmetry the three operators H, $\mathcal{M}_D = \rho_c\mathcal{M}$, and $\mathcal{M}_z$ commute with each other. For this reason we can consider the set of equations

$$H\psi = \mathcal{M}\psi$$

$$\mathcal{M}_D\psi = k\hbar\psi$$

$$\mathcal{M}_z\psi = \left(m + \frac{1}{2}\right)\hbar\psi \qquad (3.12)$$

The last two equations are closely connected with the equations for spherical harmonics with spin, which we examined in Part III.

Here we have designated the integer proportional to an eigenvalue of $\mathcal{M}_D$ by the same letter k as the integer proportional to an eigenvalue of $\mathcal{M}$ in Pauli's theory [formula (1.22), Part III]. This should not arouse confusion because in both cases k assumes the same values and the physical meaning of $\mathcal{M}$ and $\mathcal{M}_D$ in the corresponding theories is similar.

4. Generalized spherical harmonics

To find the simultaneous eigenfunctions of the operators $\mathcal{M}_D = \rho_c\mathcal{M}$ and $\mathcal{M}_z$ we must transform them to spherical coordinates. At the same time we will perform a canonical transformation of the four component wave function similar to the one that we used in Section 2, Part III, for the two component wave function of Pauli's theory.

We will denote the 4×4 matrices σ_x, σ_y, σ_z that correspond to our choice of the Dirac matrices as s_1, s_2, s_3.[2] According to formu-

---

[2] There is no reason to fear confusion of these matrices with the operators s_x, s_y, s_z, introduced in Part IV, for the components of the spin angular momentum of a system of electrons.

las (4.26), Chapter I, we have

$$s_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{bmatrix}, \quad s_2 = \begin{bmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{bmatrix}, \quad s_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$(4.1)$$

To express the operators $\mathscr{M}_z$ and $\mathscr{M}$ in spherical coordinates we can use the formulas derived in Section 2, Part III, on the basis of Pauli's theory, the only difference being that we must change the matrices σ_1, σ_2, σ_3 to s_1, s_2, s_3. Let us write the most important of these formulas using the new notation.

In the spherical coordinates r, θ, φ, which are connected with the rectangular coordinates x, y, z by the relationships

$$x = r \sin \theta \cos \varphi, \qquad y = r \sin \theta \sin \varphi, \qquad z = r \cos \theta \quad (4.2)$$

the operators $\mathscr{M}_z$ and $\mathscr{M}$ have the form

$$\mathscr{M}_z = p_\varphi + \frac{\hbar}{2} s_3 \tag{4.3}$$

$$\mathscr{M} = (- s_1 \sin \varphi + s_2 \cos \varphi) p_\theta$$
$$+ (- s_1 \cot \theta \cos \varphi - s_2 \cot \theta \sin \varphi + s_3) p_\varphi + \hbar \tag{4.4}$$

where, as usual, p_r, p_θ, p_φ stand for the operators

$$p_r = - i\hbar \frac{\partial}{\partial r}, \qquad p_\theta = i\hbar \frac{\partial}{\partial \theta}, \qquad p_\varphi = - i\hbar \frac{\partial}{\partial \varphi} \tag{4.5}$$

(the operator p_r does not enter into the expressions for $\mathscr{M}_z$ and $\mathscr{M}$).

We perform the canonical transformation of the operators and functions according to formulas similar to (6.7) and (6.8), Part III, namely

$$\mathscr{L}' = S \mathscr{L} S^+, \qquad \psi' = S \psi \tag{4.6}$$

where

$$S = \cos \frac{\varphi}{2} + i s_3 \sin \frac{\varphi}{2}, \qquad S^+ = \cos \frac{\varphi}{2} - i s_3 \sin \frac{\varphi}{2} \tag{4.7}$$

After transformation we have

$$\mathscr{M}'_z = p_\varphi \tag{4.8}$$

$$\mathscr{M}' = - s_1 \cot \theta \, p_\varphi + s_2 \left(p_\theta - \frac{i\hbar}{2} \cot \theta \right) + s_3 p_\varphi + \frac{\hbar}{2} \tag{4.9}$$

Then we perform a transformation similar to (6.21), Part III, namely

$$\mathscr{M}'' = T \mathscr{M}' T^+, \qquad \psi'' = T \psi' \tag{4.10}$$

21*

where

$$T = \cos \frac{\theta}{2} + is_2 \sin \frac{\theta}{2}, \qquad T^+ = \cos \frac{\theta}{2} - is_2 \sin \frac{\theta}{2} \qquad (4.11)$$

As in Pauli s theory we get

$$\mathscr{M}''_z = \mathscr{M}'_z = p_\varphi$$

$$\mathscr{M}'' = -\frac{s_1}{\sin \theta} p_\varphi + s_2 \left(p_\theta - \frac{i\hbar}{2} \cot \theta \right) \qquad (4.12)$$

Finally, putting

$$\mathscr{M}^* = (\sin \theta)^{1/2} \mathscr{M}'' (\sin \theta)^{-1/2}, \quad \psi^* = (\sin \theta)^{1/2} \psi'' \qquad (4.13)$$

we have

$$\mathscr{M}^* = -\frac{s_1}{\sin \theta} p_\varphi + s_2 p_\theta, \qquad \mathscr{M}^*_z = p_\varphi \qquad (4.14)$$

According to (4.7) and (4.11) the matrices of the canonical transformation, S and T, contain the operators s_1, s_2, s_3 but do not contain ρ_a, ρ_b, ρ_c. For this reason the appearance of the latter operators does not change in the transformation. Specifically, according to (4.27), Chapter I, we have

$$\rho_c^* = \rho_c = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} \qquad (4.15)$$

After multiplying by ρ_c the eigenvalue equation for $\mathscr{M}^*_D = \rho_c \mathscr{M}^*$ we can write it as

$$\mathscr{M}^* \psi^* = k\hbar \rho_c \psi^* \qquad (4.16)$$

The system of equations for the four components of ψ^*, which corresponds to (4.16), will be

$$-\frac{p_\varphi}{\sin \theta} \psi_2^* - ip_\theta \psi_2^* = -k\hbar \psi_4^*$$

$$-\frac{p_\varphi}{\sin \theta} \psi_1^* + ip_\theta \psi_1^* = k\hbar \psi_3^*$$

$$\frac{p_\varphi}{\sin \theta} \psi_4^* - ip_\theta \psi_4^* = k\hbar \psi_2^*$$

$$\frac{p_\varphi}{\sin \theta} \psi_3^* + ip_\theta \psi_3^* = -k\hbar \psi_1^* \qquad (4.17)$$

If we express the operators p_φ and p_θ in terms of derivatives and change the sign on both sides of some of these equations, we can write them as two like systems of equations for two functions

each, namely

$$\frac{i}{\sin\theta}\frac{\partial\psi_3^*}{\partial\varphi} - \frac{\partial\psi_3^*}{\partial\theta} = k\psi_1$$

$$\frac{i}{\sin\theta}\frac{\partial\psi_1^*}{\partial\varphi} + \frac{\partial\psi_1^*}{\partial\theta} = k\psi_3 \qquad (4.18)$$

and

$$\frac{i}{\sin\theta}\frac{\partial\psi_2^*}{\partial\varphi} - \frac{\partial\psi_2^*}{\partial\theta} = -k\psi_4$$

$$\frac{i}{\sin\theta}\frac{\partial\psi_4^*}{\partial\varphi} + \frac{\partial\psi_4^*}{\partial\theta} = -k\psi_2 \qquad (4.18^*)$$

Equations (4.18) and (4.18*) differ only in the sign before k.

We met these equations in Pauli's theory when we examined spherical harmonics with spin [Section 3, Part III]. We wrote them in the form

$$\frac{1}{\sin\theta}\frac{\partial Z}{\partial\varphi} - \frac{\partial Z}{\partial\theta} = kY$$

$$\frac{1}{\sin\theta}\frac{\partial Y}{\partial\varphi} + \frac{\partial Y}{\partial\theta} = kZ \qquad (4.19)$$

The solutions of our equations (4.18) and (4.18*) will then be

$$\psi_1^* = f(r)\,Y(\theta,\varphi)$$

$$\psi_2^* = g(r)\,Z(\theta,\varphi)$$

$$\psi_3^* = f(r)\,Z(\theta,\varphi)$$

$$\psi_4^* = -g(r)\,Y(\theta,\varphi) \qquad (4.20)$$

where the functions $f(r)$ and $g(r)$ no longer depend on θ and φ. Their dependence on r is determined by the eigenvalue equation for the Hamiltonian.

5. The radial equation

Now let us turn to the Hamiltonian. After going over to spherical coordinates the Hamiltonian can be written as

$$H^* = c\rho_a P^* + mc^2\rho_c + U(r) \qquad (5.1)$$

Operator P^* for four component functions is obtained from the corresponding operator for two component functions by substituting the 4×4 matrices s_1, s_2, s_3 for the 2×2 Pauli matrices σ_1, σ_2, σ_3. On the basis of (6.37), Part III, we have

$$P^* = \frac{s_1}{r}\,p_\theta + \frac{s_2}{r\sin\theta}\,p_\varphi + s_3 p_r \qquad (5.2)$$

This operator is connected with the operator

$$\mathcal{M}^* = -\frac{s_1}{\sin\theta}\,p_\varphi + s_2 p_\theta \qquad (5.3)$$

which was examined in Section 4, by the same relationship as in Pauli's theory [see (6.38), Part III], namely

$$P^* = s_3\left(p_r + i\,\frac{\mathcal{M}^*}{r}\right) \qquad (5.4)$$

We assume that the four component function ψ^* is an eigenfunction of the operator

$$\mathcal{M}_D^* = \rho_c\mathcal{M}^* \qquad (5.5)$$

which (unlike $\mathcal{M}^*$) commutes with the Hamiltonian. For this reason we can use formula (4.16) and put

$$\mathcal{M}^*\psi^* = k\hbar\rho_c\psi^* \qquad (5.6)$$

In view of (5.4) we have

$$P^*\psi^* = s_3\left(p_r + \frac{ik\hbar}{r}\,\rho_e\right)\psi^* \qquad (5.7)$$

and consequently

$$\rho_a P^*\psi^* = s_3\left(\rho_a p_r + \rho_b\,\frac{k\hbar}{r}\right)\psi^* \qquad (5.8)$$

Hence the eigenvalue equation for the Hamiltonian can be written as

$$H^*\psi^* = \left(c\rho_a s_3 p_r + c\rho_b s_3\frac{k\hbar}{r} + mc^2\rho_c + U(r)\right)\psi^* = W\psi^* \qquad (5.9)$$

This equation includes the matrices

$$\tau_a = \rho_a s_3, \qquad \tau_b = \rho_b s_3, \qquad \tau_c = \rho_c \qquad (5.10)$$

which satisfy the same relationships as ρ_a, ρ_b, ρ_c:

$$\tau_a\tau_b = i\tau_c, \qquad \tau_b\tau_c = i\tau_a, \qquad \tau_c\tau_a = i\tau_b \qquad (5.11)$$

To simplify further calculations we write the matrices τ_a, τ_b, τ_c in explicit form. We have

$$\tau_a = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \quad \tau_b = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix}, \quad \tau_c = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}$$

$$(5.12)$$

We rewrite Eq. (5.9) in the form

$$H^*\psi^* = \left(c\tau_a p_r + c\tau_b\frac{k\hbar}{r} + mc^2\tau_c + U(r)\right)\psi^* = W\psi^* \qquad (5.13)$$

Using the expressions (5.12) for the matrices τ_a, τ_b, τ_c, we can write Eq. (5.13) in the expanded form. After shifting the term with potential energy to the right side we get

$$cp_2\psi_1^* - ic\frac{k\hbar}{r}\psi_4^* - mc^2\psi_4^* = (W - U)\psi_1^*$$

$$-cp_2\psi_2^* - ic\frac{k\hbar}{r}\psi_3^* + mc^2\psi_3^* = (W - U)\psi_2^*$$

$$cp_2\psi_3^* + ic\frac{k\hbar}{r}\psi_2^* + mc^2\psi_2^* = (W - U)\psi_3^*$$

$$-cp_2\psi_4^* + ic\frac{k\hbar}{r}\psi_1^* - mc^2\psi_1^* = (W - U)\psi_4^* \qquad (5.14)$$

Now we substitute the ψ_i^* from (4.20) and replace operator p_r with the corresponding derivative. For the radial functions $f(r)$ and $g(r)$ we get the following set of equations:

$$-ic\hbar\frac{df}{dr} + ic\frac{k\hbar}{r}g + mc^2 g = (W - U)f$$

$$ic\hbar\frac{dg}{dr} - ic\frac{k\hbar}{r}f + mc^2 f = (W - U)g \qquad (5.15)$$

repeated twice.

6. Comparison with the Schrödinger equation

We can rid ourselves of the complex coefficients in (5.15) by putting

$$\frac{f + g}{\sqrt{2}} = f_1, \qquad \frac{f - g}{i\sqrt{2}} = f_2 \qquad (6.1)$$

whence

$$f_1 + if_2 = \sqrt{2}f, \qquad f_1 - if_2 = \sqrt{2}g \qquad (6.2)$$

Adding and subtracting the two equations in (5.15), we get for the new functions f_1 and f_2 a system of two equations of the first order with real coefficients:

$$\frac{df_1}{dr} - \frac{k}{r}f_1 = \frac{-mc^2 - W + U}{\hbar c}f_2$$

$$\frac{df_2}{dr} + \frac{k}{r}f_2 = \frac{-mc^2 + W - U}{\hbar c}f_1 \qquad (6.3)$$

When the energy W is close to $+mc^2$, the coefficient of f_2 in the first equation is much greater than the coefficient of f_1 in the second. For this reason f_2 is extremely small compared with f_1:

$$|f_2| \ll |f_1| \qquad (6.4)$$

Consequently, the functions f and g in (5.15) are nearly equal (and nearly real). To compare the system of equations (6.3) with

the Schrödinger equation we put

$$W = mc^2 + E \tag{6.5}$$

and consider the values of $E/(mc^2)$ and $(E - U)/(mc^2)$ extremely small compared with unity. If we ignore them, we get

$$\frac{df_1^0}{dr} - \frac{k}{r} f_1^0 = -\frac{2mc}{\hbar} f_2^0$$

$$\frac{df_2^0}{dr} + \frac{k}{r} f_2^0 = \frac{E - U}{\hbar c} f_1^0 \tag{6.6}$$

where f_1^0 and f_2^0 are approximate values of f_1 and f_2. When we exclude f_2^0 from these equations, we get the following equation for f_1^0:

$$-\frac{\hbar^2}{2m} \frac{d^2 f_1^0}{dr^2} + \frac{\hbar^2 (k-1) k}{2mr^2} f_1^0 + U f_1^0 = E f_1^0 \tag{6.7}$$

If here we put

$$f_1^0 = r R(r) \tag{6.8}$$

then the equation for $R(r)$

$$\frac{d^2 R}{dr^2} + \frac{2}{r} \frac{dR}{dr} - \frac{k(k-1)}{r^2} R + \frac{2m}{\hbar^2} [E - U(r)] R = 0 \tag{6.9}$$

coincides with the Schrödinger equation for the radial function [see (3.16), Chapter IV, Part II], provided that the Schrödinger quantum number l is linked with our quantum number k by the relationship

$$k(k-1) = l(l+1) \tag{6.10}$$

which coincides with (3.17), Part III. Thus the number l introduced in Section 3, Part III (that is, the degree of the ordinary spherical harmonics in terms of which spherical harmonics with spin are expressed) is nothing but the azimuthal quantum number of Schrödinger's theory.

We can exclude f_2 from Eqs. (6.3) without any omissions. We then get

$$\frac{d^2 f_1}{dr^2} - \frac{k(k-1)}{r^2} f_1 + \frac{2m}{\hbar^2} (E - U) f_1$$

$$= -\frac{1}{2mc^2 + E - U} \frac{dU}{dr} \left(\frac{df_1}{dr} - \frac{k}{r} f_1 \right) - \frac{(E - U)^2}{\hbar^2 c^2} f_1 \tag{6.11}$$

On the right-hand side there are small terms that represent a correction for the theory of relativity and for spin. For the two values

$$k = l + 1 \quad \text{and} \quad k = -l \tag{6.12}$$

for which the left-hand side of (6.11) is the same, the values of this correction are different. The difference of the corrections for

the energy levels (the difference of the diagonal elements of the matrix for the correction terms) gives the approximate value of the distance between the spectral terms, namely

$$\Delta E = E(k) - E(-k+1) = \frac{\hbar^2}{4m^2c^2}(2k-1)\int_0^\infty \frac{1}{r}\frac{dU}{dr}[f_1^0(r)]^2\,dr \quad (6.13)$$

where $f_1^0(r)$ is the solution to Eq. (6.7) with the following normalization:

$$\int_0^\infty [f_1^0(r)]^2\,dr = 1 \quad (6.14)$$

Here let us note one transformation of (6.11). If we put

$$f_1 = \left(1 + \frac{E-U}{2mc^2}\right)^{1/2}\varphi \quad (6.15)$$

the equation for φ will be

$$\frac{d^2\varphi}{dr^2} - \frac{k(k-1)}{r^2}\varphi + \frac{2m}{\hbar^2}(E-U)\varphi$$
$$= \frac{1}{2mc^2+E-U}\left(\frac{1}{2}\frac{d^2U}{dr^2} + \frac{k}{r}\frac{dU}{dr}\right)\varphi - \frac{(E-U)^2}{\hbar^2c^2}\varphi$$
$$+ \frac{3}{4(2mc^2+E-U)^2}\left(\frac{dU}{dr}\right)^2\varphi \quad (6.16)$$

This equation no longer contains the first derivative of the unknown function. If we assume that $|r(dU/dr)| \ll mc^2$, the last term on the right-hand of (6.16) can be dropped. In the first term on the right-hand side we can ignore $E-U$, which is much smaller than $2mc^2$.

7. General investigation of the radial equations

We want to study Eq. (6.3). These two equations have two singular points:

$$r = 0 \quad \text{and} \quad r = \infty$$

We begin by examining the region close to $r = 0$. Let us assume that for small values of r the potential energy $U(r)$ can be expanded in a power series:

$$U(r) = -\frac{A_1}{r} + A' + A''r + \ldots \quad (7.1)$$

where the expansion coefficient $-A_1$ is equal, as we noted in Section 7, Chapter IV, Part II, to the product of the charge of the nucleus, Ze, by the electron charge, $-e$, so that $A_1 = Ze^2$.

In the coefficients of the right-hand sides of Eqs. (6.3) we drop all terms but those that become infinite at $r = 0$ and we get

$$\frac{df_1}{dr} - \frac{k}{r} f_1 = - \frac{A_1}{\hbar c} \frac{1}{r} f_2 + \cdots$$

$$\frac{df_2}{dr} + \frac{k}{r} f_2 = \frac{A_1}{\hbar c} \frac{1}{r} f_1 + \cdots \qquad (7.2)$$

We assume that near $r = 0$

$$f_1 = a_1 r^\varepsilon + a_1' r^{\varepsilon+1} + \cdots$$

$$f_2 = a_2 r^\varepsilon + a_2' r^{\varepsilon+1} + \cdots \qquad (7.3)$$

and we insert these expansions into Eq. (7.2). Equating the coefficients of $r^{\varepsilon-1}$, we get a system of homogeneous linear equations

$$a_1 (\varepsilon - k) + \frac{A_1}{\hbar c} a_2 = 0$$

$$- \frac{A_1}{\hbar c} a_1 + (\varepsilon + k) a_2 = 0 \qquad (7.4)$$

to determine a_1 and a_2. These equations have a solution if the system determinant vanishes:

$$\varepsilon^2 - k^2 + \frac{A_1^2}{\hbar^2 c^2} = 0 \qquad (7.5)$$

From this for exponent ε we get two values

$$\varepsilon = \pm \left[k^2 - \frac{A_1^2}{\hbar^2 c^2} \right]^{1/2} = \pm \varepsilon_0 \qquad (7.6)$$

where ε_0 is a positive quantity:

$$\varepsilon_0 = \left[k^2 - Z^2 \left(\frac{e^2}{\hbar c} \right)^2 \right]^{1/2} \qquad (7.6^*)$$

We know that $e^2/\hbar c$ is a dimensionless quantity approximately equal to 1/137. For this reason for all possible values of Z and k the quantity in parentheses will be positive. The constant

$$\gamma = \frac{e^2}{\hbar c} = \frac{1}{137} \qquad (7.7)$$

is called the fine-structure constant.

Hence near $r = 0$ the general solution to Eqs. (6.3) has the form

$$f_1 = a r^{\varepsilon_0} (1 + \cdots) + b r^{-\varepsilon_0} \left[\frac{A_1}{\hbar c (\varepsilon_0 + k)} + \cdots \right]$$

$$f_2 = a r^{\varepsilon_0} \left[\frac{A_1}{\hbar c (\varepsilon_0 + k)} + \cdots \right] + b r^{-\varepsilon_0} (1 + \cdots) \qquad (7.8)$$

For the functions f_1 and f_2 to vanish at $r = 0$ the constant b must be zero.

If $k^2 = 1$, then $\varepsilon_0 = [1 - (Z/137)^2]^{1/2} < 1$. For this reason, though f_1 and f_2 and, consequently, ψ^* as well will vanish at $r = 0$, the initial function $\psi = \psi^*/(r \sin^{1/2}\theta)$ will become (for $|k| = 1$) infinite as $r \to 0$ and it will grow as

$$r^{\varepsilon_0 - 1} = r^{[1 - (Z/137)^2]^{1/2} - 1} \tag{7.9}$$

This constitutes a certain defect in the theory. It may arise from the fact that the Coulomb law of attraction cannot be extrapolated to distances so small that

$$\frac{Ze^2}{r} > mc^2 \tag{7.10}$$

that is,

$$r < Z \times 3 \times 10^{-13} \, \text{cm} \tag{7.11}$$

Now let us investigate the equations for large values of r. We assume as in Section 7, Chapter IV, Part II, that the potential energy far away from the nucleus is

$$U(r) = -\frac{A}{r} + \frac{B}{r^2} + \dots \tag{7.12}$$

We ask for the solution to (6.3) in the form

$$f_1 = e^{\alpha r} (a_1 r^\beta + b_1 r^{\beta - 1} + \dots)$$
$$f_2 = e^{\alpha r} (a_2 r^\beta + b_2 r^{\beta - 1} + \dots) \tag{7.13}$$

We substitute these expansions into the equations and equalize the coefficients of the terms of order $e^{\alpha r} r^\beta$ and $e^{\alpha r} r^{\beta - 1}$. We get

$$a_1 \alpha + a_2 \frac{mc^2 + W}{\hbar c} = 0$$

$$a_1 \frac{mc^2 - W}{\hbar c} + a_2 \alpha = 0 \tag{7.14}$$

and

$$b_1 \alpha + b_2 \frac{mc^2 + W}{\hbar c} = -a_1 (\beta - k) - a_2 \frac{A}{\hbar c}$$

$$b_1 \frac{mc^2 - W}{\hbar c} + b_2 \alpha = a_1 \frac{A}{\hbar c} - a_2 (\beta + k) \tag{7.15}$$

If we nullify the system determinant of (7.14), we get the following values for α:

$$\alpha = \pm \frac{1}{\hbar c} (m^2 c^4 - W^2)^{1/2} \tag{7.16}$$

The left-hand sides of (7.15) have the same coefficients as Eqs. (7.14). Since the determinant constructed from these coefficients is zero, we can exclude b_1 and b_2 from (7.15) if we multiply the first equation by $-\alpha$ and the second by $(mc^2 + W)/(\hbar c)$ and add the results. This yields

$$a_1\alpha(\beta - k) + a_2\alpha\frac{A}{\hbar c} + a_1\frac{mc^2 + W}{\hbar c}\frac{A}{\hbar c} - a_2\frac{mc^2 + W}{\hbar c}(\beta + k) = 0$$

whence, expressing $a_2\alpha$ and $a_2(mc^2 + W)/(\hbar c)$ in terms of a_1 by means of Eqs. (7.14), we get after some manipulating

$$2a_1\alpha\beta + 2a_1\frac{W}{\hbar c}\frac{A}{\hbar c} = 0$$

For β this equation gives the value

$$\beta = -\frac{AW}{\alpha\hbar^2 c^2} \tag{7.17}$$

We will not determine the constants b_1 and b_2.

In conformity with the two signs of α the general integral will have the following form:

$$f_1 = C_1\frac{mc^2 + W}{\hbar c}e^{\alpha r}r^\beta\left(1 + \frac{a_1'}{r} + \dots\right)$$
$$+ C_2\frac{mc^2 + W}{\hbar c}e^{-\alpha r}r^{-\beta}\left(1 + \frac{b_1'}{r} + \dots\right) \tag{7.18}$$

or

$$f_2 = -C_1\alpha e^{\alpha r}r^\beta\left(1 + \frac{a_2'}{r} + \dots\right)$$
$$+ C_2\alpha e^{-\alpha r}r^{-\beta}\left(1 + \frac{b_2'}{r} + \dots\right) \tag{7.18*}$$

If we assume that

$$|W| > mc^2 \tag{7.19}$$

then α and β will be pure imaginary and the functions f_1 and f_2 will remain finite as $r \to \infty$ with any choice of the constants C_1 and C_2. But we can select these constants so that f_1 and f_2 will vanish at $r = 0$. We can say, therefore, that the region (7.19) belongs to the continuous spectrum. There can be no discrete spectrum in this region because if α is pure imaginary the functions f_1 and f_2 are not square integrable.

But if

$$-mc^2 < W < +mc^2 \tag{7.20}$$

then α will be real (we will consider it positive). For this reason f_1 and f_2 will either rapidly increase (if $C_1 \neq 0$) or rapidly

fall off (if $C_1 = 0$) at infinity, so that in region (7.20) there can be no continuous spectrum. There is either a discrete spectrum or no eigenvalues at all.

Finally, if

$$W = \pm\, mc^2 \tag{7.21}$$

then α vanishes and β becomes infinite, so that (7.18) and (7.18*) break down. We must look for the asymptotic solution of our equations in a form similar to (7.11), Chapter IV, Part II. If we put

$$\alpha_0 = (8mA/\hbar^2)^{1/2} \tag{7.22}$$

we get for the case when $W = -mc^2$ by reasoning in a similar way

$$f_1 = -\,C_1\,\frac{\hbar\alpha_0}{4mc}\,e^{\alpha_0\sqrt{r}}r^{-1/4} + \ldots + C_2\,\frac{\hbar\alpha_0}{4mc}\,e^{-\alpha_0\sqrt{r}}r^{-1/4} + \ldots$$

$$f_2 = C_1 e^{\alpha_0\sqrt{r}}r^{1/4} + \ldots + C_2 e^{-\alpha_0\sqrt{r}}r^{1/4} + \ldots \tag{7.23}$$

and for the case when $W = +mc^2$

$$f_1 = C_1 e^{i\alpha_0\sqrt{r}}r^{1/4} + \ldots + C_2 e^{-i\alpha_0\sqrt{r}}r^{1/4} + \ldots$$

$$f_2 = -\,C_1\,\frac{i\hbar\alpha_0}{4mc}\,e^{i\alpha_0\sqrt{r}}r^{-1/4} + \ldots + C_2\,\frac{i\hbar\alpha_0}{4mc}\,e^{-i\alpha_0\sqrt{r}}r^{-1/4} + \ldots \tag{7.24}$$

When at great distances from the nucleus there is only attraction, $A > 0$ and α_0 is real. In this case the value $W = +mc^2$ belongs to the continuous energy spectrum and $W = -mc^2$ does not. But in the case of repulsive forces $A < 0$ and α_0 is pure imaginary; then $W = -mc^2$ belongs to the continuous spectrum and $W = +mc^2$ does not.

Hence we have established that in the case of attraction the continuous spectrum will be

$$W < -\,mc^2, \qquad W \geqslant +\,mc^2, \qquad A > 0 \tag{7.25}$$

and in the case of repulsion

$$W \leqslant -\,mc^2, \qquad W > +\,mc^2, \qquad A < 0 \tag{7.25*}$$

whereas a discrete spectrum is possible only if

$$|\,W\,| < mc^2 \tag{7.26}$$

From Eqs. (6.3) for radial functions we can derive some general corollaries regarding the distribution of the energy levels in the discrete spectrum. If we multiply the first equation in (6.3) by f_2 and the second by f_1 and add the results, we get

$$\frac{d}{dr}\,(f_1 f_2) = -\,\frac{1}{\hbar c}\,[(mc^2 + W - U)\,f_2^2 + (mc^2 - W + U)\,f_1^2] \tag{7.27}$$

By integrating this expression from 0 to ∞ and bearing in mind the behaviour of the functions of the discrete spectrum we get zero on the left, and the right-hand side yields

$$\int_0^\infty U\left(f_1^2 - f_2^2\right) dr = -\int_0^\infty \left[(mc^2 + W) f_2^2 + (mc^2 - W) f_1^2\right] dr \quad (7.28)$$

But we know that in the discrete spectrum W lies between $-mc^2$ and $+mc^2$. Whence the right-hand side is negative, and we have the inequality

$$\int_0^\infty U\left(f_1^2 - f_2^2\right) dr < 0 \quad (7.29)$$

It follows from this that at a negative U (attraction) f_1^2 is on the average greater than f_2^2. As we saw in (6.4), this is the case when W is close to $+mc^2$, that is, when $W > 0$. Consequently, in the case of attraction there are no negative energy levels belonging to the discrete spectrum.

But in the case of repulsion $(U > 0)$ there are no positive energy levels but there may be negative levels. These negative levels can have no direct physical meaning. (The same was true of the states with negative kinetic energy mentioned in Section 12, Chapter I.)

8 Quantum numbers

According to the results of our investigation the stationary state of an electron in a central field can be characterized by the energy parameter and the quantum numbers k and m, the first of which is connected with the total angular momentum and the second with its projection on the z axis. For the discrete spectrum the energy W will depend on a third (principal or radial) quantum number, which is introduced when solving the radial equation, and also on the number k, which enters this equation as a parameter. Hence here, as in Schrödinger's theory, the electronic state in the discrete spectrum is described by three quantum numbers, and the energy depends on two of them.

As we found in Section 6, the leading terms of the second-order equation analogous to the Schrödinger equation contain the quadratic expression $k(k-1)$, whereas the number k alone enters only into the correction term. The second-order equation contains $k(k-1)$ in the same way that the Schrödinger equation contains $l(l+1)$. Hence we can put

$$k(k-1) = l(l+1) \quad (8.1)$$

whence

$$\left| k - \frac{1}{2} \right| = l + \frac{1}{2} \tag{8.2}$$

For this reason the two energy levels that correspond to the same principal quantum number n and to the same l (or $|k - 1/2|$) but to two different values of k:

$$k = l + 1, \qquad k = -l \tag{8.3}$$

are quite close to each other and form a doublet. The exception is when $l = 0$. Since k cannot take on a zero value, only one level remains, and that is with $k = +1$.

The doublet separation was calculated in Section 6 [see (6.13)].

Thus Dirac's theory gives us what is required by experiment — the doubling (compared with Schrödinger's theory) of energy levels; and the level with $l = 0$ is nondegenerate as experiments require. This doubling expresses one of the two additional (intrinsic) degrees of freedom of the electron, which were mentioned in Section 12, Chapter I.

It is customary to distinguish between the two levels of a doublet by the values of a new quantum number, which we denote by j. The quantum number j, like l, is uniquely determined by k, namely

$$j = |k| - \frac{1}{2} \tag{8.4}$$

Hence j can take on positive values that are equal to an integer and a half.

Since the number of values of the magnetic quantum number m for a given k is $2|k|$, the value of j gives the multiplicity of the level, which is

$$2|k| = 2j + 1 \tag{8.5}$$

It follows from a comparison of (8.4) with (8.2) that j differs from l by $\pm 1/2$, namely

$$j = l + \frac{1}{2} \quad \text{at} \quad k > 0$$

$$j = l - \frac{1}{2} \quad \text{at} \quad k < 0 \tag{8.6}$$

Going by this formula, if we know l and j, we can find the sign of k and hence k itself.

Spectroscopy usually denotes spectral terms with different values of l, that is, $l = 0, 1, 2, \ldots,$ by the letters $S, P, D, \ldots,$ and the value of j is given as the lower index to these letters. We can relate different quantum numbers to different spectral terms in

the form of the following table:

$$k = +1, \quad l = 0, \quad j = \frac{1}{2}, \quad S$$
$$k = -1, \quad l = 1, \quad j = \frac{1}{2}, \quad P_{\frac{1}{2}}$$
$$k = +2, \quad l = 1, \quad j = \frac{3}{2}, \quad P_{\frac{3}{2}}$$
$$k = -2, \quad l = 2, \quad j = \frac{3}{2}, \quad D_{\frac{3}{2}}$$
$$k = +3, \quad l = 2, \quad j = \frac{5}{2}, \quad D_{\frac{5}{2}}$$

. . . .

It is the selection rule that determines between which spectral terms transitions are possible.

9. Heisenberg's matrices and the selection rule

We will denote the wave function corresponding to the quantum numbers n, k, m (where n is the principal quantum number) by ψ or another letter without a prime, and the wave function with the quantum numbers n', k', m' by ψ' or the alternative letter with a prime (here we discard the asterisk, which in Section 4 was used to distinguish wave functions in spherical coordinates).

If the wave function is normalized so that

$$\int \int \int \bar{\psi} \psi \, dr \, d\theta \, d\varphi = 1 \tag{9.1}$$

any element of Heisenberg's matrix for any one of the coordinates, say, x, will be

$$(nkm \mid x \mid n'k'm') = \int \int \int x \bar{\psi} \psi' \, dr \, d\theta \, d\varphi$$
$$= \int \int \int x (\bar{\psi}_1 \psi'_1 + \bar{\psi}_2 \psi'_2 + \bar{\psi}_3 \psi'_3 + \bar{\psi}_4 \psi'_4) \, dr \, d\theta \, d\varphi \tag{9.2}$$

If we insert in (9.1) the expressions for $\psi_1, \psi_2, \psi_3, \psi_4$, (4.20), the normalization condition will be

$$\int \int \int (\bar{f}f + \bar{g}g)(\bar{Y}Y + \bar{Z}Z) \, dr \, d\theta \, d\varphi = 1 \tag{9.3}$$

This holds if

$$\int_0^\infty (\bar{f}f + \bar{g}g) \, dr = 1 \tag{9.4}$$

and

$$\int_{\theta=0}^\pi \int_{\varphi=0}^{2\pi} (\bar{Y}Y + \bar{Z}Z) \, d\theta \, d\varphi = 1 \tag{9.5}$$

Now if we insert (4.20) into (9.2), we get

$$(nkm \mid x \mid n'k'm') = \int \int \int x\,(\bar{f}f' + \bar{g}g')\,(\overline{Y}Y' + \overline{Z}Z')\,dr\,d\theta\,d\varphi \quad (9.6)$$

Similar expressions can be written for the other coordinates, y and z.

As in Schrödinger's theory, the triple integrals of type (9.6) separate into the products of simple integrals. And if we put

$$r_D\,(nk;\,n'k') = \int\limits_0^\infty r\,(\bar{f}f' + \bar{g}g')\,dr \qquad (9.7)$$

then the elements of Heisenberg's matrices for the coordinates x, y, z will be equal to the product of (9.7) multiplied respectively by

$$(km \mid \sin\theta\cos\varphi \mid k'm') = \int \int \sin\theta\cos\varphi\,(\overline{Y}Y' + \overline{Z}Z')\,d\theta\,d\varphi \quad (9.8)$$

$$(km \mid \sin\theta\sin\varphi \mid k'm') = \int \int \sin\theta\sin\varphi\,(\overline{Y}Y' + \overline{Z}Z')\,d\theta\,d\varphi \quad (9.9)$$

$$(km \mid \cos\theta \mid k'm') = \int \int \cos\theta\,(\overline{Y}Y' + \overline{Z}Z')\,d\theta\,d\varphi \qquad (9.10)$$

To evaluate these integrals we express Y and Z in terms of y_1 and y_2 by means of (3.4) and (3.13), Part III. We get

$$\overline{Y}Y' + \overline{Z}Z' = \frac{1}{4\pi}\,e^{i\,(m'-m)\,\varphi}\,(y_1 y_1' + y_2 y_2')\sin\theta \qquad (9.11)$$

We first evaluate the integral (9.10). It will obviously differ from zero only if $m' = m$; in this case it will be

$$(km \mid \cos\theta \mid k'm) = \frac{1}{2}\int\limits_0^\pi \cos\theta\,(y_1 y_1' + y_2 y_2')\sin\theta\,d\theta \qquad (9.12)$$

or, if we use the notations of Section 4, Part III,

$$(km \mid \cos\theta \mid k'm) = \frac{1}{2}\int\limits_0^\pi \cos\theta\,y\,(k,\,m,\,\theta)\,y\,(k',\,m,\,\theta)\sin\theta\,d\theta \qquad (9.12^*)$$

If we now express the product $y\,(k, m, \theta)\cos\theta$ in accordance with (4.9), Part III,

$y\,(k,\,m,\,\theta)\cos\theta$

$$= -\frac{2m+1}{4k^2-1}\,y\,(-k,\,m,\,\theta) + \frac{[(k-m)\,(k+m+1)]^{1/2}}{\mid 2k+1 \mid}\,y\,(k+1,\,m,\,\theta)$$

$$+ \frac{[(k+m)\,(k-m-1)]^{1/2}}{\mid 2k-1 \mid}\,y\,(k-1,\,m,\,\theta) \quad (9.13)$$

then on the basis of the orthogonality of functions y we conclude that the integral (9.12) is nonzero in only three cases:

$$k' = -k, \qquad k' = k+1, \qquad k' = k-1 \qquad (9.14)$$

In these cases it is equal to the respective coefficients in (9.13), namely

$$(km \,|\, \cos\theta \,|\, -k\,m) = -\frac{2m+1}{4k^2-1} \qquad (9.15)$$

$$(km \,|\, \cos\theta \,|\, k+1\,m) = \frac{[(k-m)(k+m+1)]^{1/2}}{|\,2k+1\,|} \qquad (9.15^*)$$

$$(km \,|\, \cos\theta \,|\, k-1\,m) = \frac{[(k+m)(k-m-1)]^{1/2}}{|\,2k-1\,|} \qquad (9.15^{**})$$

We compute the first two integrals, (9.8) and (9.9), in a similar way. For this it is convenient to do as we did in Section 9, Chapter IV, Part II, that is, construct their linear combination

$$(km \,|\, \sin\theta\, e^{i\varphi} \,|\, k'm')$$
$$= \frac{1}{4\pi} \int\int e^{i\,(m'-m+1)\,\varphi} \sin^2\theta\, y\,(k,\,m,\,\theta)\, y\,(k',m',\,\theta')\, d\theta\, d\varphi \quad (9.16)$$

which, obviously, will differ from zero only if $m' = m - 1$. With this condition it is

$$km \,|\, \sin\theta\, e^{i\varphi} \,|\, k'\, m - 1)$$
$$= \frac{1}{2} \int_0^\pi \sin^2\theta\, y\,(k,\,m,\,\theta)\, y\,(k',\,m-1,\,\theta)\, d\theta \qquad (9.17)$$

If we now express the product $y(k', m-1, \theta) \sin\theta$ by means of (4.10), Part III,

$$y\,(k',\,m-1,\,\theta)\sin\theta = 2\,\frac{(k'^2-m^2)^{1/2}}{4k'^2-1}\, y\,(-k',\,m,\,\theta)$$
$$- \frac{[(k'-m-1)(k'-m)]}{2k'-1}\, y\,(k'-1,\,m,\,\theta)$$
$$+ \frac{[(k'+m+1)(k'+m)]^{1/2}}{2k'+1}\, y\,(k'+1,\,m,\,\theta) \qquad (9.18)$$

we find that the integral in (9.17) does not vanish in only three cases (9.14), when it is

$$(-k'\,m \,|\, \sin\theta\, e^{i\varphi} \,|\, k'\, m-1) = 2\,\frac{(k'^2-m^2)^{1/2}}{4k'^2-1}$$

$$(k'-1\,m \,|\, \sin\theta\, e^{i\varphi} \,|\, k'\, m-1) = \frac{[(k'-m-1)(k'-m)]^{1/2}}{2k'-1}$$

$$(k'+1\,m \,|\, \sin\theta\, e^{i\varphi} \,|\, k'\, m-1) = \frac{[(k'+m+1)(k'+m)]^{1/2}}{2k'+1} \qquad (9.19)$$

or, if we express k' in terms of k,

$$(km|\sin\theta\, e^{i\varphi}|-k\,m-1) = 2\,\frac{(k^2-m^2)^{1/2}}{4k^2-1}$$

$$(km|\sin\theta\, e^{i\varphi}|k+1\,m-1) = -\frac{[(k-m)(k-m+1)]^{1/2}}{2k+1}$$

$$(km|\sin\theta\, e^{i\varphi}|k-1\,m-1) = \frac{[(k+m)(k+m-1)]^{1/2}}{2k+1} \qquad (9.19^*)$$

From this we get the matrix elements (9.8) and (9.9) according to formulas similar to (9.23) and (9.24), Chapter IV, Part II. We write these elements in the form of a table:

	$(km\mid\sin\theta\cos\varphi\mid k'm')$	$(km\mid\sin\theta\sin\varphi\mid k'm')$
$k'=-k$ $m'=m-1$	$\dfrac{(k^2-m^2)^{1/2}}{4k^2-1}$	$-i\,\dfrac{(k^2-m^2)^{1/2}}{4k^2-1}$
$k'=-k$ $m'=m+1$	$\dfrac{[k^2-(m+1)^2]^{1/2}}{4k^2-1}$	$i\,\dfrac{[k^2-(m+1)^2]^{1/2}}{4k^2-1}$
$k'=k+1$ $m'=m-1$	$-\dfrac{1}{2}\dfrac{[(k-m)(k-m+1)]^{1/2}}{2k+1}$	$\dfrac{i}{2}\dfrac{[(k-m)(k-m+1)]^{1/2}}{2k+1}$
$k'=k+1$ $m'=m+1$	$\dfrac{1}{2}\dfrac{[(k+m+1)(k+m+2)]^{1/2}}{2k+1}$	$\dfrac{i}{2}\dfrac{[(k+m+1)(k+m+2)]^{1/2}}{2k+1}$
$k'=k-1$ $m'=m-1$	$\dfrac{1}{2}\dfrac{[(k+m)(k+m-1)]^{1/2}}{2k-1}$	$-\dfrac{i}{2}\dfrac{[(k+m)(k+m-1)]^{1/2}}{2k-1}$
$k'=k-1$ $m'=m+1$	$-\dfrac{1}{2}\dfrac{[(k-m-1)(k-m-2)]^{1/2}}{2k-1}$	$\dfrac{i}{2}\dfrac{[(k-m-1)(k-m-2)]^{1/2}}{2k-1}$

These results reflect the selection rule that determines between which terms transitions are possible or impossible.

The selection rule with respect to quantum number m will be the same as in Schrödinger's theory; namely, for coordinate z (light polarized along the z axis)

$$m'=m \qquad (9.20)$$

and for coordinates x and y (light polarized on the xy plane)

$$m'=m\pm 1 \qquad (9.21)$$

Levels that differ from each other by the value of the quantum number m can be distinguished only in a magnetic field directed along the z axis. It is not surprising then that the z axis plays

such a special role in the selection rule with respect to m — the direction of this axis is physically selected by the direction of the magnetic field.

The selection rule with respect to k is

$$k' = -k, \qquad k' = k + 1, \qquad k' = k - 1 \qquad (9.22)$$

This rule states that quantum number $l = |k - {}^1/_2| - {}^1/_2$ always changes by unity, as in Schrödinger's theory. However, not all transitions of the type $l' = l \pm 1$ are possible. There is a second condition for quantum number j: it must either remain unchanged or change only by unity. For example, a transition is possible between the $P_{{}^1/_2}$ and $D_{{}^3/_2}$ terms but not between the $P_{{}^1/_2}$ and $D_{{}^5/_2}$ terms.

10. Alternative derivation of the selection rule

In view of the importance of the selection rule we will note another derivation, less elementary but not requiring a knowledge of spherical harmonics. The idea of this derivation belongs to Dirac.

Consider the operator

$$\mathcal{M}_z = m_z + \frac{1}{2}\hbar\sigma_z$$

with eigenvalues

$$\left(m + \frac{1}{2}\right)\hbar$$

The matrix representing this operator will be diagonal in quantum number m. If we write only this quantum number and assume the rest, we get

$$(m|\,\mathcal{M}_z\,|m') = \left(m + \frac{1}{2}\right)\hbar\delta_{mm'}$$

Now we consider the matrix for coordinate z with elements

$$(m|\,z\,|m')$$

From the commutation relation

$$\mathcal{M}_z z - z\mathcal{M}_z = 0 \qquad (10.1)$$

we get the following relationship between the matrix elements:

$$\left(m + \frac{1}{2}\right)\hbar\,(m|\,z\,|m') - (m|\,z\,|m')\left(m' + \frac{1}{2}\right)\hbar = 0 \qquad (10.2)$$

or

$$(m - m')\,(m|\,z\,|m') = 0 \qquad (10.2^*)$$

Consequently, the only matrix elements of z that differ from zero are those for which $m' = m$. As we know, this reflects the selection rule for z with respect to quantum number m.

Now we note that the selection rules for x, y, z are the same as for $\dot{x} = c\rho_a\sigma_x$, $\dot{y} = c\rho_a\sigma_y$, $\dot{z} = c\rho_a\sigma_z$. This is why instead of x, y, z we can deal with the matrices

$$\alpha_1 = \rho_a\sigma_x, \qquad \alpha_2 = \rho_a\sigma_y, \qquad \alpha_3 = \rho_a\sigma_z$$

which in some cases is simpler. For instance, from

$$\mathscr{M}_z\alpha_3 - \alpha_3\mathscr{M}_z = 0$$

or

$$\mathscr{M}_z\dot{z} - \dot{z}\mathscr{M}_z = 0 \tag{10.3}$$

it follows that

$$(m - m')(m \,|\, \dot{z} \,|\, m') = 0 \tag{10.4}$$

which is what we obtained before.

We will derive the selection rule for x and y or for $\dot{x}$ and $\dot{y}$, which is the same. We have

$$\mathscr{M}_z\dot{x} - \dot{x}\mathscr{M}_z = c\rho_a(\mathscr{M}_z\sigma_x - \sigma_x\mathscr{M}_z) = i\hbar c\rho_a\sigma_y$$

or

$$\mathscr{M}_z\dot{x} - \dot{x}\mathscr{M}_z = i\hbar\dot{y} \tag{10.5}$$

and in the same way

$$\mathscr{M}_z\dot{y} - \dot{y}\mathscr{M}_z = c\rho_a(\mathscr{M}_z\sigma_y - \sigma_y\mathscr{M}_z) = - i\hbar c\rho_a\sigma_x$$

or

$$\mathscr{M}_z\dot{y} - \dot{y}\mathscr{M}_z = - i\hbar\dot{x} \tag{10.6}$$

If we multiply (10.6) by i and add the product to (10.5), we get

$$\mathscr{M}_z(\dot{x} + i\dot{y}) - (\dot{x} + i\dot{y})\mathscr{M}_z = \hbar(\dot{x} + i\dot{y}) \tag{10.7}$$

Transferring to matrix elements we get

$$(m - m' - 1)(m \,|\, \dot{x} + i\dot{y} \,|\, m') = 0 \tag{10.8}$$

which is the result that can be obtained from (9.16).

In a similar manner we obtain

$$(m - m' + 1)(m \,|\, \dot{x} - i\dot{y} \,|\, m') = 0 \tag{10.9}$$

Whence the condition that the matrix elements for x and for y do not vanish is

$$m' = m \pm 1 \tag{10.10}$$

The way we proceeded can be somewhat altered. It follows from (10.5) and (10.6) that

$$\mathscr{M}_z^2\dot{x} - 2\mathscr{M}_z\dot{x}\mathscr{M}_z + \dot{x}\mathscr{M}_z^2 - \hbar^2\dot{x} = 0 \tag{10.11}$$

Transferring to matrix elements we get

$$\left(m + \frac{1}{2}\right)^2 (m \mid \dot{x} \mid m') - 2 \left(m + \frac{1}{2}\right) (m \mid \dot{x} \mid m') \left(m' + \frac{1}{2}\right)$$
$$+ (m \mid x' \mid m') \left(m' + \frac{1}{2}\right)^2 - (m \mid \dot{x} \mid m') = 0$$

or

$$[(m - m')^2 - 1] (m \mid \dot{x} \mid m') = 0 \tag{10.12}$$

from which we get the previous result (10.10).

We now seek the selection rule with respect to quantum number k. The quantity $k\hbar$ is the eigenvalue of the operator

$$\mathcal{M}_D = \rho_c \mathcal{M} \tag{10.13}$$

where

$$\mathcal{M} = \sigma_x m_x + \sigma_y m_y + \sigma_z m_z + \hbar \tag{10.14}$$

We considered $\mathcal{M}_D$ in Section 3 (it is the generalization of the corresponding operator in Pauli's theory). By formula (1.18), Part III, operator $\mathcal{M}$ satisfies the relationship

$$\mathcal{M}^2 = \hbar \mathcal{M} + \left(m_x^2 + m_y^2 + m_z^2\right) \tag{10.15}$$

Let us consider the operator

$$\mathcal{L} = \mathcal{M}_D \left(\mathcal{M}_D^2 \dot{z} - \dot{z} \mathcal{M}_D^2\right) - \left(\mathcal{M}_D^2 \dot{z} - \dot{z} \mathcal{M}_D^2\right) \mathcal{M}_D \tag{10.16}$$

In view of (10.13) and the equality

$$\mathcal{M}_D^2 = \mathcal{M}^2 \tag{10.17}$$

we can write $\mathcal{L}$ in the following form:

$$\mathcal{L} = (\rho_c \mathcal{M}) \left(\mathcal{M}^2 \dot{z} - \dot{z} \mathcal{M}^2\right) - \left(\mathcal{M}^2 \dot{z} - \dot{z} \mathcal{M}^2\right) (\rho_c \mathcal{M}) \tag{10.18}$$

Since matrix ρ_c commutes with $\mathcal{M}$ and anticommutes with $\dot{z} = c\rho_a\sigma_z$, we can write $\mathcal{L}$ as

$$\mathcal{L} = \rho_c \left[\mathcal{M} \left(\mathcal{M}^2 \dot{z} - \dot{z} \mathcal{M}^2\right) + \left(\mathcal{M}^2 \dot{z} - \dot{z} \mathcal{M}^2\right) \mathcal{M}\right] \tag{10.19}$$

But owing to (10.15)

$$\mathcal{M}^2 \dot{z} - \dot{z} \mathcal{M}^2 = \hbar \left(\mathcal{M} \dot{z} - \dot{z} \mathcal{M}\right) \tag{10.20}$$

which yields

$$\mathcal{L} = \hbar \rho_c \left[\mathcal{M} \left(\mathcal{M} \dot{z} - \dot{z} \mathcal{M}\right) + \left(\mathcal{M} \dot{z} - \dot{z} \mathcal{M}\right) \mathcal{M}\right] \tag{10.19*}$$

instead of (10.19). The terms of type $\mathcal{M} \dot{z} \mathcal{M}$ cancel out and we get

$$\mathcal{L} = \hbar \rho_c \left(\mathcal{M}^2 \dot{z} - \dot{z} \mathcal{M}^2\right) \tag{10.21}$$

After applying (10.20) once more we arrive at

$$\mathcal{L} = \hbar^2 \rho_c \left(\mathcal{M} \dot{z} - \dot{z} \mathcal{M}\right) \tag{10.22}$$

If we return to operator $\mathcal{M}_D$ of Dirac's theory and take into account that $\dot{z}$ and ρ_c anticommute, we can write

$$\mathscr{L} = \hbar^2 (\mathcal{M}_D \dot{z} + \dot{z}\mathcal{M}_D) \tag{10.23}$$

By equating the initial and final expressions (10.16) and (10.23) for $\mathscr{L}$, we come to the equality

$$\mathcal{M}_D^3 \dot{z} - \mathcal{M}_D \dot{z}\mathcal{M}_D^2 - \mathcal{M}_D^2 \dot{z}\mathcal{M}_D + \dot{z}\mathcal{M}_D^3 - \hbar^2 (\mathcal{M}_D \dot{z} + \dot{z}\mathcal{M}_D) = 0 \tag{10.24}$$

Now we transfer from operators to matrices and use the representation in which $\mathcal{M}_D$ is diagonal. The matrix element for each term in (10.24) is obtained from the matrix element $(k|\dot{z}|k')$ by multiplying the latter into the eigenvalues $\hbar k$ and $\hbar k'$, the eigenvalues of $\mathcal{M}_D$, in the proper degree (into $\hbar k$ if $\mathcal{M}_D$ is to the left of z and by $\hbar k'$ if $\mathcal{M}_D$ is to the right of $\dot{z}$). Dividing out $\hbar^3$, we get

$$(k^3 - kk'^2 - k^2 k' + k'^3 - k - k')(k|\dot{z}|k') = 0 \tag{10.25}$$

or

$$(k + k')(k - k' - 1)(k - k' + 1)(k|\dot{z}|k') = 0 \tag{10.25*}$$

From this follows the selection rule with respect to k:

$$k' = -k, \qquad k' = k + 1, \qquad k' = k - 1 \tag{10.26}$$

which we have already derived in a different way.

11. The hydrogen atom. Radial functions

For the hydrogen atom, for which the potential energy is

$$U(r) = -\frac{e^2}{r} \tag{11.1}$$

radial equations (6.3) permit an exact solution. In the given case these equations have the form

$$\frac{df_1}{dr} - \frac{k}{r} f_1 = \frac{1}{\hbar c} \left(-mc^2 - W - \frac{e^2}{r} \right) f_2$$

$$\frac{df_2}{dr} + \frac{k}{r} f_2 = \frac{1}{\hbar c} \left(-mc^2 + W + \frac{e^2}{r} \right) f_1 \tag{11.2}$$

We will confine ourselves to examining the discrete spectrum, that is, when $W^2 < m^2 c^4$. We put

$$\alpha = \frac{(m^2 c^4 - W^2)^{1/2}}{\hbar c} \tag{11.3}$$

and consider α positive. Having in mind the asymptotic formulas (7.18), we introduce a new independent variable

$$x = 2\alpha r \tag{11.4}$$

We also put

$$W = mc^2 \cos \varepsilon, \qquad a = \frac{mc}{\hbar} \sin \varepsilon, \qquad 0 < \varepsilon < \pi \qquad (11.5)$$

Finally, we recall the value of the fine-structure constant

$$\gamma = \frac{e^2}{\hbar c} = \frac{1}{137} \qquad (11.6)$$

After we change variables Eqs. (11.2) appear as

$$\frac{df_1}{dx} - \frac{k}{x} f_1 = \left(-\frac{1}{2} \cot \frac{\varepsilon}{2} - \frac{\gamma}{x} \right) f_2$$

$$\frac{df_2}{dx} + \frac{k}{x} f_2 = \left(-\frac{1}{2} \tan \frac{\varepsilon}{2} + \frac{\gamma}{x} \right) f_1 \qquad (11.7)$$

Angle ε is a parameter here. We must determine it in such a way that Eqs. (11.7) have solutions that are finite and continuous throughout the interval $0 < x < \infty$ and that vanish at $x = 0$ and $x = \infty$.

Let us introduce two new functions, F and G, such that

$$f_1 = \frac{F - G}{2 \sin (\varepsilon/2)}, \qquad f_2 = \frac{F + G}{2 \cos (\varepsilon/2)} \qquad (11.8)$$

The new functions are expressed in terms of f_1 and f_2 in the following way:

$$F(x) = f_1 \sin \frac{\varepsilon}{2} + f_2 \cos \frac{\varepsilon}{2}$$

$$G(x) = -f_1 \sin \frac{\varepsilon}{2} + f_2 \cos \frac{\varepsilon}{2} \qquad (11.9)$$

Multiplying the first equation in (11.7) by $\pm \sin (\varepsilon/2)$, the second by $\cos (\varepsilon/2)$, and adding, we get

$$\frac{dF}{dx} + \frac{k}{x} G = -\frac{1}{2} F + \frac{\gamma}{x \sin \varepsilon} (F \cos \varepsilon - G)$$

$$\frac{dG}{dx} + \frac{k}{x} F = \frac{1}{2} G + \frac{\gamma}{x \sin \varepsilon} (F - G \cos \varepsilon) \qquad (11.10)$$

We can exclude G or F from these equations. The result will be

$$x^2 \frac{d^2 F}{dx^2} + x \frac{dF}{dx} + \left[-\frac{1}{4} x^2 + x \left(\gamma \cot \varepsilon + \frac{1}{2} \right) - k^2 + \gamma^2 \right] F = 0 \quad (11.11)$$

or

$$x^2 \frac{d^2 G}{dx^2} + x \frac{dG}{dx} + \left[-\frac{1}{4} x^2 + x \left(\gamma \cot \varepsilon - \frac{1}{2} \right) - k^2 + \gamma^2 \right] G = 0 \quad (11.11^*)$$

These equations are of the same type as

$$-\frac{d}{dx} \left(x \frac{dy}{dx} \right) + \left(\frac{x}{4} + \frac{s^2}{4x} \right) y = \left(p + \frac{s+1}{2} \right) y \qquad (11.12)$$

which we considered in detail in the chapter devoted to a nonrelativistic electron in a Coulomb field (Sections 3, 4, and 5, Chapter V, Part II). To make Eq. (11.11) for F coincide with (11.12) for y it is sufficient to put

$$s = 2(k^2 - \gamma^2)^{1/2}, \qquad p + \frac{s}{2} = \gamma \cot \varepsilon \qquad (11.13)$$

For (11.11*) parameter s will have the same value and p will be less by one unit. Hence the eigenvalues are

$$\gamma \cot \varepsilon = p + (k^2 - \gamma^2)^{1/2}, \qquad p = 0, 1, 2, \ldots \qquad (11.14)$$

and the eigenfunctions are

$$F(x) = C x^{s/2} e^{-x/2} Q_p^s(x) \qquad (11.15)$$

$$G(x) = C' x^{s/2} e^{-x/2} Q_{p-1}^s(x) \qquad (11.15^*)$$

Since F and G are linked by the system of equations (11.10), the ratio of the constants, C/C', will be a definite number.

Solving the first of the equations in (11.10) with respect to G, we get

$$G(x) = \frac{1}{k + \gamma/\sin \varepsilon} \left[x \frac{dF}{dx} + \left(p + \frac{s}{2} - \frac{x}{2} \right) F \right]$$

$$= \frac{C}{k + \gamma/\sin \varepsilon} x^{s/2} e^{-x/2} \left(p Q_p^s - x \frac{dQ_p^s}{dx} \right) \qquad (11.16)$$

But in view of the property of polynomial Q_p^s derived earlier we have

$$p Q_p^s - x \frac{dQ_p^s}{dx} = p(p + s) Q_{p-1}^s$$

$$[(4.13^*), \text{ Chapter V, Part II}]$$

It also follows from (11.13) that

$$p(p + s) = \frac{\gamma^2}{\sin^2 \varepsilon} - k^2 \qquad (11.17)$$

For this reason

$$G(x) = C \left(\frac{\gamma}{\sin \varepsilon} - k \right) x^{s/2} e^{-x/2} Q_{p-1}^s(x) \qquad (11.18)$$

Hence we have expressed the constants in (11.15*), C', in terms of C.

To determine C we turn to the normalization condition. As we did in Schrödinger's theory, we introduce the atomic unit of length

$$a = \frac{\hbar^2}{me^2} \qquad (11.19)$$

and denote the distance from the nucleus in atomic units by

$$r_1 = \frac{r}{a} \tag{11.20}$$

Constant α introduced by (11.3), will be

$$\alpha = \frac{(m^2 c^4 - W^2)^{1/2}}{\hbar c} = \sin \varepsilon / (a\gamma) \tag{11.21}$$

so that variable x is connected with r_1 by the relationship

$$x = 2\alpha r = 2 \frac{\sin \varepsilon}{\gamma} r_1 \tag{11.22}$$

We take

$$\int_0^\infty (f_1^2 + f_2^2) \, dr_1 = 1 \tag{11.23}$$

as the normalization condition, which also means that

$$\int_0^\infty (f_1^2 + f_2^2) \, dx = \frac{2 \sin \varepsilon}{\gamma} \tag{11.23*}$$

If we express f_1 and f_2 in terms of F and G, we get

$$f_1^2 + f_2^2 = \frac{1}{\sin^2 \varepsilon} (F^2 + G^2 - 2FG \cos \varepsilon) \tag{11.24}$$

Substituting (11.24) into (11.23*) and remembering that F and G are orthogonal to each other, we can write the normalization condition as follows:

$$\int_0^\infty (F^2 + G^2) \, dx = \frac{2 \sin^3 \varepsilon}{\gamma} \tag{11.23**}$$

From this we can find the normalization constant C:

$$C^2 = \frac{1}{(p-1)! \, \Gamma (p+s)(-k+\gamma/\sin \varepsilon)} \frac{\sin^4 \varepsilon}{\gamma^2} \tag{11.25}$$

or by (11.17)

$$C^2 = \frac{k + \gamma/\sin \varepsilon}{p! \, \Gamma (p+s+1)} \frac{\sin^4 \varepsilon}{\gamma^2} \tag{11.25*}$$

From this, if we introduce

$$Q_p^{*s}(x) = \frac{1}{[p! \, \Gamma (p+s+1)]^{1/2}} Q_p^s(x)$$

we get

$$F(x) = \frac{\sin^2 \varepsilon}{\gamma} \left(\frac{\gamma}{\sin \varepsilon} + k \right)^{1/2} x^{s/2} e^{-x/2} Q_p^{*s}(x) \tag{11.26}$$

$$G(x) = \frac{\sin^2 \varepsilon}{\gamma} \left(\frac{\gamma}{\sin \varepsilon} - k \right)^{1/2} x^{s/2} e^{-x/2} Q_{p-1}^{*s}(x) \tag{11.26*}$$

It is convenient to denote $\gamma/\sin\varepsilon$ by n^*:

$$n^* = \frac{\gamma}{\sin\varepsilon} \tag{11.27}$$

Its square is

$$n^{*2} = p^2 + 2p\,(k^2 - \gamma^2)^{1/2} + k^2 \tag{11.28}$$

so that n^* differs only slightly from the integer

$$n = p + |k| \tag{11.29}$$

which can be interpreted as the principal quantum number.

By inserting n^* into (11.26) and (11.26*) we can write the two expressions as

$$F = \frac{\gamma}{n^{*2}}\,(n^* + k)^{1/2}\left(\frac{2r_1}{n^*}\right)^{s/2} e^{-r_1/n^*} Q_p^{*s}\left(\frac{2r_1}{n^*}\right) \tag{11.30}$$

$$G = \frac{\gamma}{n^{*2}}\,(n^* - k)^{1/2}\left(\frac{2r_1}{n^*}\right)^{s/2} e^{-r_1/n^*} Q_{p-1}^{*s}\left(\frac{2r_1}{n^*}\right) \tag{11.30*}$$

For a given value of the principal quantum number n the number k can take on the following sequence:

$$-n+1,\ -n+2,\ \ldots,\ -1,\ +1,\ \ldots,\ n-1,\ n \tag{11.31}$$

in all $2n-1$ values. The number k cannot be equal to $-n$ since then the lower index of Q_{p-1}^{*s} in (11.30) would become negative. But the value $k = +n$ is possible because in this case from (11.29) we get $p = 0$, and from (11.28) we get $n^* = k$, so that the factor $(n^* - k)^{1/2}$ of Q_{p-1}^{*s} vanishes.

The number p is closely connected with the radial quantum number n_r of Schrödinger's theory. Namely, because of

$$n = n_r + l + 1 = p + |k|$$

and the relationship between l and k we have

$$p = n_r \qquad \text{at} \quad k > 0$$
$$p = n_r + 1 \qquad \text{at} \quad k < 0 \tag{11.32}$$

Thus, we have found the eigenfunctions corresponding to the discrete spectrum. It presents no difficulty to solve our equations for the value $W = +mc^2$, which corresponds to the boundary between the discrete and continuous spectra, and for the continuous spectrum. But we will not do so.

12. Fine-structure levels of hydrogen

Now we will express energy in terms of quantum numbers. According to (11.5) and (11.14),

$$W = mc^2 \cos\varepsilon = \frac{mc^2\,[p + (k^2 - \gamma^2)^{1/2}]}{\{[p + (k^2 - \gamma^2)^{1/2}]^2 + \gamma^2\}^{1/2}} \tag{12.1}$$

or by means of (11.27)

$$W = mc^2 \left(1 - \frac{\gamma^2}{n^{*2}} \right)^{1/2} \tag{12.1*}$$

Formula (12.1) is called the *Sommerfeld formula.*

As we noted at the end of Section 7, Dirac's theory gives only positive levels. The lowest level (the ground state of hydrogen) corresponds to quantum numbers $k = +1$, $p = 0$, $(n = 1)$, and is

$$W_0 = mc^2 (1 - \gamma^2)^{1/2} \tag{12.2}$$

The entire discrete spectrum is located in the interval

$$mc^2 (1 - \gamma^2)^{1/2} \leqslant W < mc^2 \tag{12.3}$$

whereas in the interval

$$- mc^2 < W < mc^2 (1 - \gamma^2)^{1/2} \tag{12.4}$$

there are no energy eigenvalues at all.

To compare the Sommerfeld formula with the Bohr formula, which we derived in Schrödinger's theory, we will go on to approximate formulas. By approximately extracting the square root in (12.1*) we get

$$W = mc^2 \left(1 - \frac{\gamma^2}{n^{*2}} \right)^{1/2} = mc^2 - \frac{1}{2} \frac{mc^2 \gamma^2}{n^{*2}} - \frac{1}{8} \frac{mc^2 \gamma^4}{n^{*4}} \tag{12.5}$$

But by (11.6) and (11.19)

$$mc^2 \gamma^2 = \frac{e^2}{a} \tag{12.6}$$

whereas by (11.28) and (11.29)

$$n^{*2} = n^2 + 2 (n - | k |) [(k^2 - \gamma^2)^{1/2} - | k |] \tag{12.7}$$

Using these expressions, we get, within terms of the order of $mc^2 \gamma^6$,

$$W = mc^2 - \frac{e^2}{2an^2} - \frac{e^2}{8a} \frac{\gamma^2}{n^3} \left(\frac{4}{| k |} - \frac{3}{n} \right) \tag{12.8}$$

The first term is a constant (the relativistic rest energy), the second is the (nonrelativistic) formula for the Bohr levels of hydrogen, and the third term provides a relativistic correction to it. This correction depends upon k as well as upon the principal quantum number n. For this reason the energy levels of hydrogen, which in Schrödinger's theory depended only on n and did not change when the azimuthal quantum number l changed, split into several levels lying close to each other. These are obtained if we give the number k in (12.8) all permissible values (11.31). The result is the fine structure of the hydrogen spectral lines observed

in experiments. We note that the energy levels depend only on the absolute value of k (that is, on j but not on l), so that, say, the spectral terms $P_{3/2}$ and $D_{3/2}$ for hydrogen coincide.

Next we will find the doublet separation in the central-field problem (the doublet in alkali metals), that is, the value

$$\Delta W = W(n, k) - W(n, -k+1) \tag{12.9}$$

If we consider $k > 0$ and put $k = l + 1$, the formula (12.8) gives

$$\Delta W = \frac{e^2 \gamma^2}{2an^3 l (l+1)} \tag{12.10}$$

We computed the same quantity for the general case of a central field in Section 6 [see formula (6.13)]. Let us apply that formula to hydrogen. We have

$$\Delta E = \frac{\hbar^2}{4m^2c^2} (2l+1) \int_0^\infty \frac{e^2}{r^3} [f_1^0(r)]^2 dr \tag{12.11}$$

Here we must put

$$r = ar_1, \qquad [f_1^0(r)]^2 dr = r_1^2 R_{nl}^2(r_1) dr_1 \tag{12.12}$$

so that

$$\Delta E = \frac{e^2 \gamma^2}{4a} (2l+1) \int_0^\infty R_{nl}^2(r_1) \frac{dr_1}{r_1} \tag{12.13}$$

Using the expression (6.11), Chapter V, Part II, for $R_{nl}(r_1)$ and introducing the integration variable $x = 2r_1/n$, we get

$$\Delta E = \frac{e^2 \gamma^2}{a} \frac{2l+1}{n^4} \int_0^\infty x^{2l-1} e^{-x} [Q_{n-l-1}^{*2l+1}(x)]^2 dx \tag{12.14}$$

We denote the integral by I. If we put $n - l - 1 = p$ and $2l + 1 = s$, we can write the integral as

$$I = \int_0^\infty x^{s-2} e^{-x} [Q_p^{*s}(x)]^2 dx \tag{12.15}$$

This integral was calculated in Section 4, Chapter V, Part II [formulas (4.21) and (4.23)]. It is

$$I = \frac{2p + s + 1}{(s-1) s (s+1)} = \frac{n}{2(2l+1) l (l+1)} \tag{12.16}$$

Substitution of I in (12.14) yields

$$\Delta E = \frac{e}{2a} \frac{\gamma^2}{n^3 l (l+1)} \tag{12.17}$$

that is, the previous result (12.10).

13. The Zeeman effect. Statement of the problem

The energy levels for a central field, as we have seen, depend only on the quantum numbers n and k. The third quantum number m does not enter into the fine structure formula (12.8), so that one level can correspond to states with different values of m.

But if we place the atom in a magnetic field, each level splits into several levels that differ by the value of m. This is what is called the *Zeeman effect*. Schrödinger's theory proves insufficient to explain this effect. But Dirac's theory, as we will now see, provides a full explanation of this effect and one that corresponds to experimental data in every respect.

In Section 5, Part III, we examined a generalization of the Schrödinger equation for the case of a magnetic field, namely, the Pauli equation. But the Pauli equation disregards relativistic corrections. At the same time the splitting of levels in a magnetic field is, generally speaking, of the same order of magnitude as these corrections. Hence both must be considered simultaneously. For this reason to explain the Zeeman effect we must examine the Dirac equation, which takes into account both the theory of relativity and the magnetic field.

We assume that we have a constant magnetic field $\mathscr{H}$ directed along the z axis. As we know, the vector potential of this field will be

$$A_x = -\frac{1}{2} |\mathscr{H}| y, \qquad A_y = \frac{1}{2} |\mathscr{H}| x, \qquad A_z = 0 \qquad (13.1)$$

and the generalized component of the vector potential corresponding to angle φ can be found by the formula

$$A_x \, dx + A_y \, dy = A_\varphi \, d\varphi$$

or

$$\frac{1}{2} |\mathscr{H}| (x \, dy - y \, dx) = \frac{1}{2} |\mathscr{H}| \rho^2 \, d\varphi$$

whence

$$A_\varphi = \frac{1}{2} |\mathscr{H}| \rho^2 = \frac{1}{2} |\mathscr{H}| r^2 \sin^2 \theta \qquad (13.2)$$

The additional term we must add to the Hamiltonian in the absence of a field, or the zeroth-order Hamiltonian

$$H^* = c\rho_a \left(\frac{s_1}{r} p_\theta + \frac{s_2}{r \sin \theta} p_\varphi + s_3 p_r \right) + mc^2 \rho_c + U(r) \qquad (13.3)$$

is found according to the general rule by replacing p_φ with

$$P_\varphi = p_\varphi + \frac{e}{c} A_\varphi$$

We denote this term by R. It will be

$$R = \frac{1}{2} e |\, \mathcal{H} \,| \rho_a s_2 r \sin \theta \qquad (13.4)$$

To find the correction energy due to this term we must apply perturbation theory and calculate the matrix elements of R that correspond to the various transitions. As we saw in Section 7, Chapter II, Part II, the main role is played by the matrix elements that correspond to the transitions between levels that are very close to each other, that is, between the levels of one doublet in our case. Hence we will consider none but these transitions and put

$$n' = n \qquad (13.5)$$

$$k'(k'-1) = k(k-1) \qquad (13.5^*)$$

As for the quantum number m, operator R commutes with the operator of differentiation with respect to φ, that is, with $\mathcal{M}_z$, whose eigenvalues are $(m + \frac{1}{2})\hbar$. For this reason (see the second derivation of the selection rule in Section 10) the matrix for R will be diagonal in m, so that we must put

$$m' = m \qquad (13.6)$$

Therefore we ask for the following matrix elements:

$$(k\,|\,R\,|\,k), \qquad (k\,|\,R\,|-k+1)$$
$$(-k+1\,|\,R\,|\,k), \qquad (-k+1\,|\,R\,|-k+1) \qquad (13.7)$$

(for brevity we omit the quantum numbers n and m).
The total Hamiltonian is

$$H = H^* + R$$

and the eigenvalue equation will be

$$(H^* + R)\,\psi = W\psi \qquad (13.8)$$

Using the method outlined in Section 7, Chapter II, Part II, we seek the approximate eigenfunction as

$$\psi = c_1 \psi_k^* + c_2 \psi_{-k+1}^* \qquad (13.9)$$

where ψ^* are the eigenfunctions of the zeroth-order Hamiltonian, H^*. We substitute (13.9) into (13.8), premultiply by $\overline{\psi_k^*}$ and integrate. We also multiply the result of the above substitution into $\overline{\psi_{-k+1}^*}$ and integrate. This gives us two equations

$$[W_k + (k\,|\,R\,|\,k)]\,c_1 + (k\,|\,R\,|-k+1)\,c_2 = Wc_1$$
$$(-k+1\,|\,R\,|\,k)\,c_1 + [W_{-k+1} + (-k+1\,|\,R\,|-k+1)]\,c_2 = Wc_2 \qquad (13.10)$$

with W_k the unperturbed energy level corresponding to a given value of k. We nullify the system determinant of (13.10), which is constructed from the coefficients of the unknowns c_1 and c_2. This yields for W a quadratic equation whose roots are

$$W = \frac{1}{2} (W_k + W_{-k+1} + R_k + R_{-k+1})$$

$$\pm \frac{1}{2} [(W_k - W_{-k+1} + R_k - R_{-k+1})^2 + 4 | (- k + 1 | R | k) |^2]^{1/2} \quad (13.11)$$

where for brevity we put

$$R_k = (k | R | k) \quad (13.12)$$

Equation (13.11) gives the corrected values of the energy level.

14. Calculation of the perturbation matrix

Now we turn to computing the matrix elements (13.7). Using the value of $\alpha_2 = \rho_a s_2$ given by (4.29), Chapter I, and the expression (4.20) for the function ψ^*, we get

$$(ek | R | k')$$

$$= \frac{1}{2} e | \mathcal{H} | \iiint r \sin \theta \, i(\bar{g}f' - \bar{f}g') (\bar{Z}Y' - \bar{Y}Z') \, dr \, d\theta \, d\varphi \quad (14.1)$$

Let us express f and g in terms of f_1 and f_2 with the help of (6.2):

$$f = \frac{f_1 + if_2}{\sqrt{2}}, \qquad g = \frac{f_1 - if_2}{\sqrt{2}} \quad (14.2)$$

and consider f_1 and f_2 real. As for the functions $Y(\theta, \varphi)$ and $Z(\theta, \varphi)$ of θ and φ, we express them as

$$Y(\theta, \varphi) = \frac{1}{(4\pi)^{1/2}} e^{i(m + 1/2)\varphi} A(\theta)$$

$$Z(\theta, \varphi) = \frac{1}{(4\pi)^{1/2}} e^{i(m + 1/2)\varphi} B(\theta) \quad (14.3)$$

in terms of functions $A(\theta)$ and $B(\theta)$, which depend only on one angle, θ [see Section 3, Part III]. Since at $m = m'$ the integrand in (14.1) does not depend on φ, integration with respect to φ amounts to multiplying by 2π. We get

$$(k | R | k') = -\frac{1}{2} e | \mathcal{H} | \int_0^\infty r (f_1 f_2' + f_2 f_1') \, dr \int_0^\infty \sin \theta (BA' + AB') \, d\theta \quad (14.4)$$

We will show that the left integral can be found approximately without solving the radial equations. According to (6.3) these

equations have the form

$$\frac{df_1}{dr} - \frac{k}{2} f_1 = \frac{-mc^2 - W + U}{\hbar c} f_2$$

$$\frac{df_2}{dr} + \frac{k}{2} f_2 = \frac{-mc^2 + W - U}{\hbar c} f_1 \qquad (14.5)$$

In Section 7 we saw that the radial functions of a discrete spectrum fall off at infinity exponentially and vanish at $r = 0$. For this reason we have the identity

$$\int_0^\infty (f_1 f_1' + f_2 f_2') \, dr = -\int_0^\infty r \frac{d}{dr} (f_1 f_1' + f_2 f_2') \, dr$$

$$= -\int_0^\infty r \left(f_1 \frac{df_1'}{dr} + f_1' \frac{df_1}{dr} + f_2 \frac{df_2'}{dr} + f_2' \frac{df_2}{dr} \right) dr \qquad (14.6)$$

On the right-hand side we replace the derivatives with their expressions given by differential equations (14.5) and we get

$$\int_0^\infty (f_1 f_1' + f_2 f_2') \, dr = -(k + k') \int_0^\infty (f_1 f_1' - f_2 f_2') \, dr$$

$$+ \frac{2mc}{\hbar} \int_0^\infty r (f_1 f_2' + f_2 f_1') \, dr$$

$$+ \frac{W' - W}{\hbar c} \int_0^\infty r (f_1 f_2' - f_1' f_2) \, dr \qquad (14.7)$$

The second term on the right-hand side is the sought integral that enters into (14.4). We note now that f_2 is very small compared with f_1 and that for the values of k and k' we are considering, for which $k(k-1) = k'(k'-1)$, the difference $W' - W$ (the doublet separation) is very small compared with $2mc^2$, and f_1' differs very little from f_1 [both approximately satisfy the same equation (6.7)]. If we take into account the normalization of the functions, we get from (14.7) the approximate formula for the integral with respect to r in (14.4):

$$\int_0^\infty r (f_1 f_2' + f_2 f_1') \, dr = \frac{\hbar}{2mc} (k + k' + 1) \qquad (14.8)$$

If we introduce the *Larmor frequency*

$$\omega_L = \frac{e |\mathcal{H}|}{2mc} \qquad (14.9)$$

and insert (14.8) into (14.4), we get

$$(k \mid R \mid k') = -\tfrac{1}{4} \hbar \omega_L \, (k + k' + 1) \int_0^\pi \sin \theta \, (BA' + AB') \, d\theta \quad (14.10)$$

Now we have only to compute the integral

$$J = \int_0^\pi \sin \theta \, (BA' + AB') \, d\theta \qquad (14.11)$$

Integrating by parts and using Eq. (3.5), Part III, for the functions A and B, we get

$$J = \int_0^\pi \cos \theta \, \frac{d}{d\theta} \, (BA' + AB') \, d\theta$$

$$= -(k + k') \int_0^\pi \cos \theta \, (AA' - BB') \, d\theta \qquad (14.12)$$

We multiply (14.11) by $k + k'$ and add the result to (14.12). This yields

$$(k + k' + 1) J = -(k + k') \, \mathrm{Re} \left(\int_0^\pi e^{i\theta} (A + iB)(A' + iB') \, d\theta \right) \quad (14.13)$$

Now according to the formula

$$A + iB = (\sin \theta)^{1/2} \, e^{-i\theta/2} \, (y_1 + iy_2) \qquad (14.14)$$

we introduce the functions y_1 and y_2, used earlier in Pauli's theory [see (3.13), Part III]. Then

$$(k + k' + 1) J = -(k + k') \int_0^\pi (y_1 y_1' - y_2 y_2') \sin \theta \, d\theta \quad (14.13^*)$$

so that

$$(k \mid R \mid k') = \tfrac{1}{4} \hbar \omega_L \, (k + k') \int_0^\pi (y_1 y_1' - y_2 y_2') \sin \theta \, d\theta \quad (14.15)$$

Owing to the orthogonality of the spherical harmonics P_l^{*m}, which can be used to express y_1 and y_2, the integral (14.15) differs from zero only when $l' = l$, that is, when (13.5) holds. Hence the matrix elements (13.7) are not only the most important for calculating corrections but the only ones that differ from zero (if $n' = n$).

To evaluate the integral (14.15) it is sufficient to express y_1 and y_2 according to (3.24), Part III, in terms of ordinary spheri-

cal harmonics and the normalization condition for the latter. As a result we get

$$(k \mid R \mid k') = \frac{1}{4} \hbar \omega_L \frac{k + k'}{|2k - 1|} \left\{ \frac{(k + m)(k' + m)}{[|(k + m)(k' + m)|]^{1/2}} \right.$$
$$\left. - [|(k - m - 1)(k' - m - 1)|]^{1/2} \right\} \quad (14.16)$$

Giving k' the values k and $-k + 1$, we obtain

$$R_k = (k \mid R \mid k) = \hbar \omega_L \frac{k}{k - 1/2} \left(m + \frac{1}{2} \right) \quad (14.17)$$

$$(k \mid R \mid -k + 1) = - \hbar \omega_L \frac{[(k + m)(k - m - 1)]^{1/2}}{|2k - 1|} \quad (14.18)$$

Replacing k by $-k + 1$ in (14.17) yields

$$R_{-k+1} = (-k + 1 \mid R \mid -k + 1) = \hbar \omega_L \frac{k - 1}{k - 1/2} \left(m + \frac{1}{2} \right) \quad (14.19)$$

We have now computed all the matrix elements in (13.7).

15. Splitting of energy levels in a magnetic field

To find the shifted energy levels we need only insert the expressions found for the matrix elements of R into formula (13.11). For brevity we denote the half-sum of the spectral terms in a doublet by W_0:

$$\frac{1}{2} (W_k + W_{-k+1}) = W_0 \quad (15.1)$$

and put

$$W_k - W_{-k+1} = \Delta W \quad (15.2$$

We insert these expressions along with (14.17), (14.18), and (14.19) into (13.11) and get

$$W = W_0 + \hbar \omega_L (m + 1/2)$$
$$\pm \frac{1}{2} \left((\Delta W)^2 + 2 \Delta W \hbar \omega_L \frac{m + 1/2}{k - 1/2} + \hbar^2 \omega_L^2 \right)^{1/2} \quad (15.3)$$

This formula fully describes the Zeeman effect.

When the magnetic field is weak, so that $\hbar \omega_L$ is small compared with the doublet separation ΔW, we can extract the square root approximately by dropping $\hbar^2 \omega_L^2$. The result will be two levels:

$$W' = W_0 + \frac{1}{2} \Delta W + \hbar \omega_L \left(m + \frac{1}{2} \right) \frac{k}{k - 1/2} \quad (15.4)$$

$$W'' = W_0 - \frac{1}{2} \Delta W + \hbar \omega_L \left(m + \frac{1}{2} \right) \frac{k - 1}{k - 1/2} \quad (15.4^*)$$

or, if instead of W_0 and ΔW we insert their values (15.1) and (15.2),

$$W' = W_k + \hbar\omega_L \left(m + \frac{1}{2} \right) \frac{k}{k - 1/2} \tag{15.5}$$

$$W'' = W_{-k+1} + \hbar\omega_L \left(m + \frac{1}{2} \right) \frac{k-1}{k - 1/2} \tag{15.5*}$$

In a magnetic field each spectral term splits into $2|k|$ separate terms corresponding to

$$m = -|k|, \ -|k| + 1, \ \ldots, \ |k| - 1 \tag{15.6}$$

The distance between two adjacent terms is

$$\hbar\omega_L \frac{k}{k - 1/2} = \hbar\omega_L g \tag{15.7}$$

where

$$g = \frac{k}{k - 1/2} \tag{15.8}$$

is what is known as the *Landé factor* (also called the *g*-factor). Since it is always positive, it can be represented as

$$g = \frac{|k|}{|k - 1/2|} = \frac{j + 1/2}{l + 1/2} \tag{15.9}$$

where j and l are defined as usual. For different spectral terms, corresponding to $k = 1, -1, 2, -2, \ldots$, the Landé factor takes on the following values:

k	l	j	term	g
1	0	$1/2$	S	2
−1	1	$1/2$	$P_{1/2}$	$2/3$
2	1	$3/2$	$P_{3/2}$	$4/3$
−2	2	$3/2$	$D_{3/2}$	$4/5$
3	2	$5/2$	$D_{5/2}$	$6/5$
−3	3	$5/2$	$F_{5/2}$	$6/7$

The case just considered is referred to as the *anomalous Zeeman effect.*

We shall now turn to the *normal Zeeman effect*, to be found in strong magnetic fields. When a field is so strong that $\hbar\omega_L$ is great compared with ΔW, we can extract the square root approximately by ignoring the square of ΔW. We then get two levels

$$W^* = W_0 + \hbar\omega_L (m + 1) + \Delta W \frac{m + 1/2}{2k - 1} \tag{15.10}$$

$$W^{**} = W_0 + \hbar\omega_L m - \Delta W \frac{m + 1/2}{2k - 1} \tag{15.11}$$

or, if we ignore ΔW as well,

$$W^* = W_0 + \hbar\omega_L\,(m + 1) \tag{15.12}$$

$$W^{**} = W_0 + \hbar\omega_L m \tag{15.13}$$

In this case the distance between the components of the *Zeeman multiplet* no longer depends on quantum number k and is equal to $\hbar\omega_L$. Thus, when the magnetic field gets stronger, several components corresponding to one value of m but to different k's merge to form one. This is the transition from the anomalous Zeeman effect to the normal Zeeman effect.

If $\Delta W > 0$, then when the field gets stronger the spectral term W' becomes W^* and the term W'' becomes W^{**}. But if $\Delta W < 0$, then W' becomes W^{**} and W'' becomes W^*. Since the square root in (15.3) keeps its sign when $\hbar\omega_L$ changes, the two spectral terms do not intersect when the magnetic field changes.

Experimental data confirm this in every detail. In fact, the formulas derived here were first found empirically.

The Zeeman effect makes it possible to compare with experiment the relative intensities of the lines that correspond to the transitions between spectral terms with given k and k' and different values of m and m'. These intensities can be calculated without a knowledge of radial functions. Indeed, in expressions of the type (9.6) for elements of Heisenberg's matrices the factor $r_D(nk; n'k')$ does not depend on m or m'. For this reason, according to (3.16), Chapter III, Part II, the intensities will be proportional to the quantities

$$I = [\,|\,(km\,|\sin\theta\cos\varphi|\,k'm')\,|^2$$
$$+\,|\,(km\,|\sin\theta\sin\varphi|\,k'm')\,|^2 + |\,(km\,|\cos\theta|\,k'm')\,|^2] \tag{15.14}$$

Using the table in Section 9 and the formulas (9.15), (9.15*), and (9.15**), for example we get for the value $k' = k + 1$ and for the cases $m' = m - 1$, $m' = m$, and $m' = m + 1$ the following values of I:

$$I_- = \frac{1}{2}\,\frac{(k - m)\,(k - m + 1)}{(2k + 1)^2}, \qquad m' = m - 1 \tag{15.15}$$

$$I_0 = \frac{(k - m)\,(k + m + 1)}{(2k + 1)^2}, \qquad m' = m \tag{15.16}$$

$$I_+ = \frac{1}{2}\,\frac{(k + m + 1)\,(k + m + 2)}{(2k + 1)^2}, \qquad m' = m + 1 \tag{15.17}$$

Here I_0 gives the intensity of light polarized in the direction of the magnetic field, and I_- and I_+ give the intensity of light polarized in the plane perpendicular to this direction. The sum of

these quantities

$$l_- + I_0 + I_+ = \frac{k+1}{2k+1}$$

does not depend on m.

We note that the factor $r_D(nk; n'k')$ primarily depends only on the quantum numbers l and l' and is approximately equal to the corresponding quantity $r(nl; n'l')$ in Schrödinger's theory [see (9.14), Chapter IV, Part II], so that its value is almost the same for the two components of a doublet. This makes it possible to compare the intensities of the Zeeman components belonging to different components of the doublet.

Chapter III

ON THE THEORY OF POSITRONS

1. Charge conjugation

In Chapter I of Dirac's theory (more exactly, in Section 5) we noted that it is possible to select the matrices in such a way that the system of four equations for the components of the wave function of a free electron will have real coefficients [Eqs. (5.9)]. According to (5.8), Chapter I, the corresponding Dirac matrices will be

$$\alpha_1^0 = \sigma_1, \qquad \alpha_2^0 = \rho_2\sigma_2, \qquad \alpha_3^0 = \sigma_3, \qquad \alpha_4^0 = \rho_3\sigma_2 \qquad (1.1)$$

where σ_1, σ_2, σ_3 and ρ_1, ρ_2, ρ_3 are the Dirac matrices (3.18) and (3.19), Chapter I. The matrix elements of α_1^0, α_2^0, α_3^0 are real, and the matrix elements of α_4^0 are pure imaginary.

Let us write the Dirac equation for an electron in an electromagnetic field that corresponds to this choice of matrices. We have

$$H\psi^0 = i\hbar\,\frac{\partial\psi^0}{\partial t} \qquad (1.2)$$

For stationary states, when the wave function depends on time via the factor $e^{-iWt/\hbar}$, we get

$$H\psi^0 = W\psi^0 \qquad (1.3)$$

where

$$H\psi^0 = \alpha_1^0\left(-i\hbar c\,\frac{\partial\psi^0}{\partial x} + eA_x\psi^0\right) + \alpha_2^0\left(-i\hbar c\,\frac{\partial\psi^0}{\partial y} + eA_y\psi^0\right)$$

$$+ \alpha_3^0\left(-i\hbar c\,\frac{\partial\psi^0}{\partial z} + eA_z\psi^0\right) + mc^2\alpha_4^0\psi^0 - e\Phi\psi^0 = W\psi^0 \qquad (1.4)$$

Now we write the equations that are complex conjugate to (1.4), and let us change the sign on both sides. We get

$$\alpha_1^0\left(-i\hbar c\,\frac{\partial\overline{\psi^0}}{\partial x} - eA_x\overline{\psi_0}\right) + \alpha_2^0\left(-i\hbar c\,\frac{\partial\overline{\psi^0}}{\partial y} - eA_y\overline{\psi^0}\right)$$

$$+ \alpha_3^0\left(-i\hbar c\,\frac{\partial\overline{\psi^0}}{\partial z} - eA_z\overline{\psi^0}\right) + mc^2\alpha_4^0\overline{\psi^0} + e\Phi\overline{\psi^0} = -W\overline{\psi^0} \qquad (1.5)$$

These equations differ from the previous (aside from the substitution of $\overline{\psi^0}$ for ψ^0) only in the different signs of the electron charge

—*e* and the energy *W*. Hence the quantity that is complex conjugate to the wave function of a particle with a negative charge and negative energy (that is, to the wave function of an electron with negative energy) can in a certain sense be associated with the wave function of a particle that has a positive charge and positive energy (that is, with the wave function of a positron in a state with positive energy). This is not a direct association, however, and cannot be interpreted in the context of the one-body problem. Such an interpretation would conflict with the fundamental concepts of quantum mechanics. But it does clear the way for an interpretation of the mysterious states of the electron with negative kinetic energy and the related second internal degree of freedom of the electron, which we noted in Section 12, Chapter I.

2. Basic ideas of positron theory

The physical interpretation of the second internal degree of freedom of the electron is based on the study of the problem of a physical system that consists of a variable number of charged particles but in which the total charge does not change (the law of conservation of the total charge).

The mathematical statement of this problem requires introducing operators of a new kind, namely, operators that act not on the wave function of a fixed number of variables (corresponding to the number of particles) but on a sequence of wave functions that depend on the variables for one, two, three, etc. particles. These new operators convert each of these functions into a function of variables for a number of particles greater by one (the *creation operator*) or less by one (the *annihilation operator*).

These operators of a new kind are a formal generalization of the one-particle wave function. The creation operators generalize the wave function itself, and the annihilation operators generalize the quantity complex conjugate to it.

Operators that represent a formal generalization of the wave function are sometimes referred to as the *quantized wave function*, and the transition from an ordinary wave function to a quantized is known as *second quantization*. Second quantization is also applied to systems consisting of an indefinite number of uncharged particles — light quanta, or photons. This is the subject of quantum electrodynamics.

The theory of second quantization will not be considered in this book. It can be found in the author's *Studies in Quantum Field Theory*. [3]

---

[3] V. A. Fock, *Studies in Quantum Field Theory*, Leningrad University Press, Leningrad, 1957 (in Russian). (The book contains articles that were first published in 1928-37.)

3. Positrons as unfilled states

In conclusion, a few words about the model proposed by Dirac for positrons as unfilled states (*vacancies*, or *holes*) in some fictitious "sea" (sometimes called the *Dirac sea*) of electrons in states with negative kinetic (and total) energy. The electric charge of this "sea" is somehow neutralized.

In a certain sense the model of a "sea" of filled negative-energy states is a generalization (or extrapolation) of the concept of a filled electron shell in an atom, whose negative charge is neutralized by the positive charge of the atomic nucleus. The lack of one electron in the atomic shell is expressed in the existence of one positive charge equal in absolute value to the electron charge and constitutes a kind of analogue or model of the positron.

However, application of the concept of a "sea" to the theory of positrons has logical faults. First, it is not understandable why this "sea" proves to be neutralized. Hence the initial concept of a "neutralized sea" is nonphysical. Second, the concept of a full set of states (which the Dirac sea is supposed to be) has mathematical meaning only when these states are discrete. In the case of a continuous spectrum a full set of states cannot be distinguished from a partially filled set, and the initial concept of the theory loses mathematical meaning.

Therefore we must recognize that there is still no closed, logically faultless theory of positrons. Construction of such a theory will probably require essentially new physical concepts in addition to those used in ordinary quantum mechanics.

For this reason this book did not set out to give a full picture of the present state of the theory of positrons but confined itself to general remarks about the fundamentals of the theory.

INDEX

absolutization of a physical process
14
absorption spectrum 146, 147
acceleration operator 96, 97
action function 76, 77
  for mass point 79
action integral 77, 227
adjoint operator (*see* hermitian conjugate operator)
Airy functions 155
Airy's differential equation 155
allowed transitions 146, 185
alpha-decay 156-159
angular momentum 161, 301-303
  commutation relations for 162
  conservation of 161-163
  in spherical coordinates 164
  spin-orbit scalar of 236
  total spin 258
angular momentum operator(s) 54, 55
  eigenfunctions and eigenvalues of 55, 56
  total 235-238
angular momentum vector 56
anharmonic oscillator, one-dimensional 133-136
  perturbation Hamiltonian for 133
annihilation operator 360
anomalous Zeeman effect 355, 356
anticommutation relations 233
anticommutator 246
antisymmetric wave function 256
area integral 161fn
associated Legendre polynomials 169, 172, 173
associate operator (*see* hermitian conjugate operator)
asymptotic series 206
atomic core 270
atomic object, experiments with 17
  interaction with an instrument 18, 19
atomic phenomena 13
atomic system, quantum description 51
atomic units 188, 189
azimuthal quantum number 180

Balmer series 199
Bessel function 200
Bohr, N. 5, 18, 117, 271
Bohr magneton 248, 311
Bohr's correspondence principle 43, 117

Bohr's frequency condition, and time derivative of operators 71
Bohr's frequency relation 143
Bohr's principles 22
Boltzmann statistics 153
boundary conditions 30
  homogeneous 30

canonically conjugate variables 42, 43
canonical transformation 59, 63, 80, 289
  and contact transformation 80-84
  and harmonic oscillator 106-109
  operator of 64, 65
  and unitary transformation 66
Cayley-Klein parameters
charge conjugation 359
charge density 138, 139
classical mechanics, applicability of 115
  Hamilton-Jacobi equation 75, 76
closed electron shell 272
commutativity of operators 52, 54, 92
  and simultaneous eigenfunctions 53
commuting operators 27
complementary properties 18
complete beta function 203
completeness 37, 39, 40, 92
Compton scattering 277
confluent hypergeometric function 192
conjugate operator (*see* hermitian conjugate operator)
conservation of energy, in quantum mechanics 75
constants of the motion 98, 99, 320
  angular momentum 161, 162
  in Dirac's theory 303
  square of angular momentum 163
contact transformation 42, 80
  and canonical transformation 80-84
continuous spectrum, for attractive forces 333
  intensities in 146-148
  for repulsive forces 333
correspondence principle 43, 117
Coulomb field 160, 188-231, 278
creation operator 360
current density 138, 139
cylindrical coordinates 217

de Broglie, L. 16, 74, 75, 93, 231
de Broglie's wave 230, 231
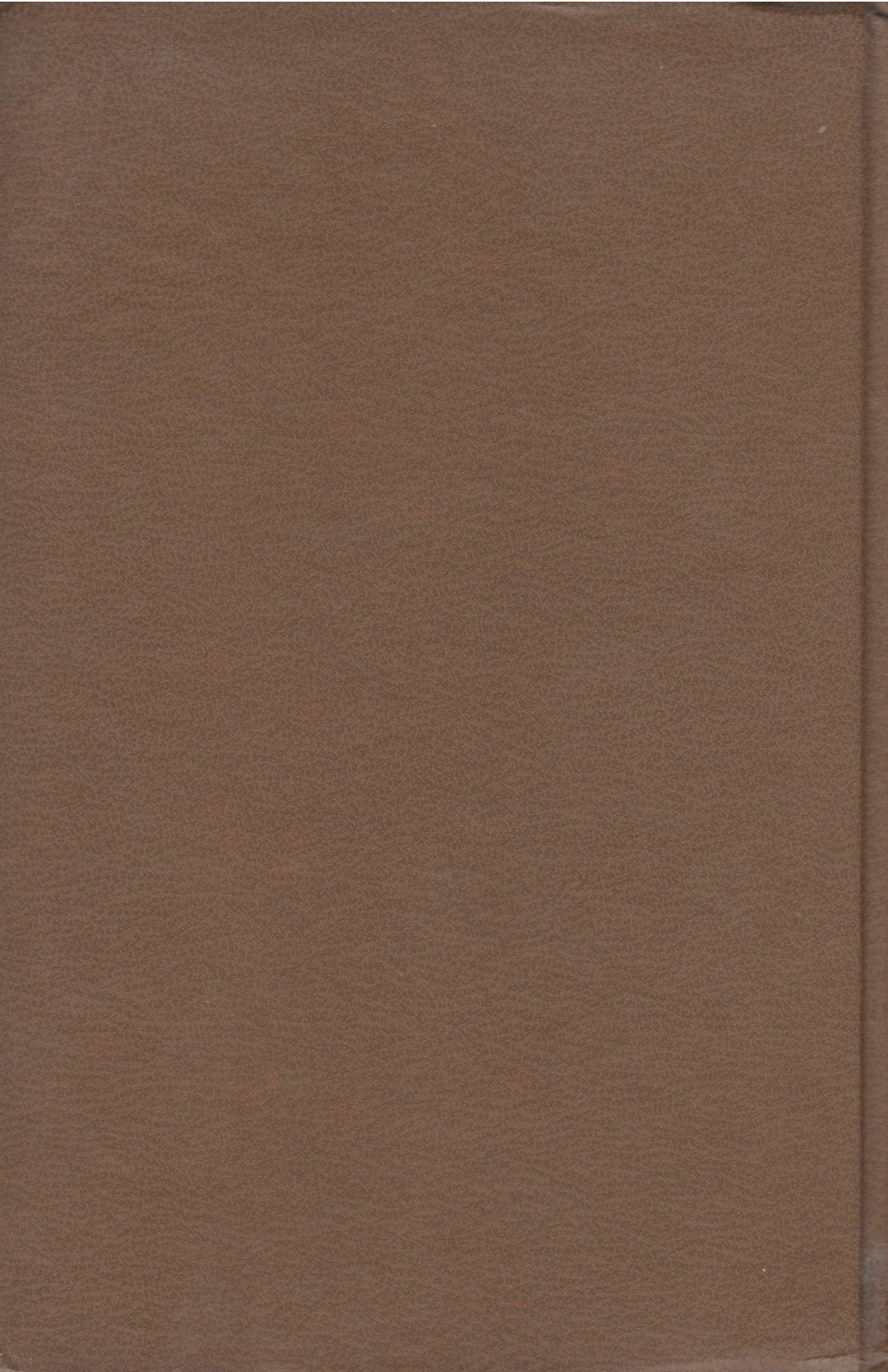de Broglie wavelength 116, 117

TO THE READER

Mir Publishers would be grateful for
your comments on the content, translation
and design of this book.
We would also be pleased to receive
any other suggestions
you may wish to make.

Our address is:
Mir Publishers
2 Pervy Rizhsky Pereulok
I-110, GSP, Moscow, 129820
USSR

Vladimir A. Fock
(1898-1974)

Vladimir Aleksandrovich Fock was one of the group of brilliant physics theoreticians whose work built the magnificent edifice of the quantum theory. A contemporary of Niels Bohr, Lev Landau, Werner Heisenberg, and Paul Dirac, he contributed much to practically all fields of theoretical and mathematical physics. His books *The Theory of Space, Time and Gravitation,* and *Electromagnetic Diffraction and Propagation Problems* have been translated into English (Pergamon Press). In 1936 Vladimir Fock merited the Mendeleev prize for his work in quantum theory of atoms, in 1946 the State prize for his work in the propagation of radio waves, and in 1960 the Lenin prize for his work in quantum field theory. In 1932 Vladimir Fock became a Corresponding Member of the USSR Academy of Sciences and in 1939 a Full Member.